


Brain texture as a marker of transdiagnostic clinical profiles in patients with recent-onset psychosis and depression

Received: 24 March 2023

Accepted: 17 October 2023

Published online: 2 January 2024

 Check for updates

A list of authors and their affiliations appears at the end of the paper

The inter-relationships of voxels can be captured by the radiomics texture features across multiple spatial scales. Prediction models of brain texture changes captured by the contrast texture feature in recent-onset psychosis (ROP) and recent-onset depression (ROD) have recently been proposed, although the validation of these models transdiagnostically at the individual level and the investigation of the variability in clinical profiles are lacking. Established prevention and treatment approaches focus on specific diagnoses and do not address the heterogeneity and manifold potential outcomes of patients. Here we aimed to investigate the cross-sectional utility of brain texture changes for (1) identification of the psychopathological state (ROP and ROD) and (2) the association of individualized brain texture maps with clinical symptom severity and outcome profiles. We developed transdiagnostic models based on structural magnetic resonance imaging data for 116 patients with ROD, 122 patients with ROP and 197 healthy control participants from the PRONIA (Personalized pROgNostic tools for early psychosIs mAnagement) study by applying a set of tools and frameworks to explain the classification decisions of the deep-learning algorithm (named explainable artificial intelligence) and clustering analysis. We investigated the contrast texture feature as the key feature for the identification of a general psychopathological state. The discrimination power of the trained prediction model was >72% and was validated in a second independent age- and sex-matched sample of 137 ROP, 94 ROD and 159 healthy control participants. Clustering analysis was implemented to map the changes in texture brain produced from an explainable artificial intelligence algorithm, in a group fashion. The explained individualized brain contrast map grouped into eight homogeneous clusters. In the clinical group, we investigated the association between the explained brain contrast texture map and clinical symptom severity as well as outcome profiles. Different patterns in the explained brain contrast texture map showed unique associations of brain alterations with clinical symptom severity and clinical outcomes, that is, age, positive, negative and depressive symptoms, and functionality. In some clusters, the mean explained brain contrast

texture map values and/or brain contrast texture voxels that contributed significantly to the classification decision predicted accurately the PANSS (positive and negative symptom scale) scores, functionality and change in functionality over time. In conclusion, we created homogeneous clusters which predict the clinical severity and outcome profile in ROP and ROD patients.

In mental health, diagnosis and treatment depend nearly entirely on a syndromal classification that defines diagnostic categories based on symptom clusters^{1–3}. However, comorbidity has long been recognized as the rule rather than the exception, with half of the patients that fulfill the diagnostic criteria for one disease also meeting the criteria for at least one other condition^{4,5}.

Common etiologies and symptoms are observed in psychotic disorders and depression, which may indicate overarching, transdiagnostic pathogenetic mechanisms^{6,7}. Hence, a broader approach that acknowledges the complexity of mental illnesses is essential to unravel the determinants of specific psychopathologies^{8–10}. Such transdiagnostic approaches may be helpful in advancing therapeutic discovery by enabling symptoms to be investigated on multiple levels of analysis, for example, manifest psychiatric syndromes versus latent psychopathology dimensions. Potential advantages include the ability to classify and treat multiple psychological disorders using the same treatment protocol, as well as the capacity to better address comorbidities.

The interconnection of psychotic disorders and depressive syndromes showcases the above issue. Depressive symptoms can be observed in the prodromal stage before the first manifestation of psychotic symptoms, during the first or later psychotic episodes, after the fading out of an acute episode and as a sign of relapse⁸. Depression may result from the neurobiological processes of psychotic disorders⁹, as a side effect of antipsychotic drugs, as an outward manifestation of problematic symptoms or as a psychological response to psychosis as a significant life experience¹⁰. For an explanation of the connection between depression and psychosis, numerous models have been presented^{11,12}. Despite the growing emphasis on transdiagnostic approaches to mental health, it is unclear if distinct diagnostic categories share a common neuronal base and our knowledge is based on meta-analyses of studies that were carried out on individual disorders using different methodologies¹³.

Structural magnetic resonance imaging (sMRI) studies have identified common brain abnormalities in patients with schizophrenia and depression (for example, in cerebral white matter¹⁴ or the cerebellum^{15,16}). Volume deficits and structural brain abnormalities have been linked to both schizophrenia^{17,18} and depression^{19,20}. These changes have also been observed in patients with recent-onset psychosis (ROP)^{17,21–24} and recent-onset depression (ROD)^{25–28}, and have a wide distribution, affecting the frontal, temporal and parietal cortical regions as well as the subcortical, cerebellar, insula and callosal regions^{29–31}. Moreover, abnormalities in the white matter (WM) of the brain and cerebrospinal fluid (CSF) have been reported in both psychotic and depressive disorders^{32–35}. However, it is unclear how the observed brain changes translate to clinical symptoms and how different brain-change profiles correspond to different symptom profiles, that is, whether or not positive and negative symptoms, depressive symptoms and functionality are associated with different brain topologies.

Despite having information from structural brain imaging studies on psychosis that date back more than 40 years, so far it has not been possible to find a diagnostic or predictive biomarker for use in clinical settings. Small sample sizes, ambiguous biomarker definitions and a dearth of replications³⁶ have been suggested to be responsible for these shortcomings. On the basis of brain volumetric changes, cross-sectional studies have reported different symptom clusters (positive

and negative symptoms, depressive symptoms and functionality) that are associated with distinct neuroimaging profiles across both diagnoses^{25,37}. In the present analysis, the contrast texture feature map expresses the gray-level differences of non-segmented magnetic resonance images, capturing inter-relationships of the segments and the cortical folding. Combining the captured information with an explainable artificial intelligence algorithm, which explains the classification decision, gives insights regarding the association of the pattern that the algorithm uses for classification with the clinical profile and outcome, grouped in homogeneous clusters of the symptom. The current study seeks (1) to investigate whether or not a transdiagnostic set of structural alterations can characterize ROP and ROD, and if any such alterations may reflect markers of psychiatric illness, and (2) to investigate associations between transdiagnostically individualized brain markers and clinical symptom severity as well as outcome profiles. We used a large training sample of ROP and ROD data and an independent validation sample, both of which are from the Personalized pROgnostic tools for early psychosis management (PRONIA) study²⁷. The measures of interest were radiomics texture features, based on previous findings by our group, where the contrast texture feature map has been identified as the key feature for psychosis^{14,37}. Mathematically, the contrast feature expresses the shape abnormalities in the brain cortex and the surface relief of the cortical boundary. Finally, we performed explainable artificial intelligence analysis, a set of tools which supports the explanation of the classification decision, to highlight the brain regions that most contributed to differences between the ROD and ROP patients. We hypothesized that (1) ROP and ROD present shared contrast texture feature abnormalities and (2) different contrast feature patterns are associated with different transdiagnostic clinical profiles. The goal of our study was to investigate the usefulness of neuroimaging markers for the transdiagnostic prediction of outcome profiles. In this sense, our results serve mainly as a proof of concept that should be followed up in larger samples. We believe that there is merit in our results as they showcase the feasibility and potential relevance of transdiagnostic approaches.

Results

Sample characteristics

Sociodemographic and clinical characteristics of the three participant groups (ROP, ROD and healthy control (HC)) for the training sample are presented in Table 1. In the training sample, there were no statistically significant differences between ROP, ROD and HC with respect to age, sex and ethnicity. There were statistically significant differences between ROP and ROD with respect to specific psychopathological symptoms and psychosocial functioning (see Table 1). The power analysis on the training sample showed that 94 ROP + ROD and 78 HC is the minimum sample size that results in a comparative fit index greater than 0.9.

In the validation baseline sample (age- and sex-matched with the training sample—see ‘Training and validation samples’ in the Methods), there were no statistically significant differences between ROP, ROD and HC with respect to age, sex and ethnicity. There were statistically significant differences between ROP and ROD with respect to specific psychopathological symptoms and psychosocial functioning (see Table 1).

Table 1 | Clinical and demographic sample details for the training sample at baseline and the independent validation sample

Group	HC	ROD	ROP	Merged ROP-ROD	ANOVA between HC, ROP and ROD	Merged ROP-ROD at follow-up (after one year)
Training sample at the baseline						
Number, <i>n</i>	197	116	122	238		203
Age (years), mean (s.d.)	26.8 (6.99)	27.5 (5.98)	26.5 (5.37)	25.63 (6.14)	$F(2,432)=0.772$ ($P=0.463$)	26.94 (5.84)
Sex (<i>n</i> for male (female))	89 (108)	55 (61)	69 (53)	124 (114)	$F(2,432)=4.59$ ($P=0.011$)	111 (54.68%)
Ethnicity (%)					$F(2,210)=1.97$ ($P=0.142$)	
White	83.98	87.14	77.97	82.71		77.54
Black	2.14	—	2.77	1.33		—
Asian	10.54	2.66	10.12	6.21		11.04
Caribbean	—	1.76	1.80	1.84		—
Other	3.34	8.44	7.34	7.91		11.42
PANSS_positive (s.d.)	—	8.23 (2.18)	11.9 (6.26)	10.0 (4.99)	$F(1,701.7)=32.58$, $P=0.001^*$	10.03 (4.99)
PANSS_negative (s.d.)	—	11.6 (5.61)	14.1 (7.91)	12.8 (6.93)	$F(1,322.1)=6.91$, $P=0.009^*$	12.78 (6.92)
PANSS_excitement (s.d.)	—	11.4 (2.63)	14.8 (6.84)	13.0 (5.41)	$F(1,616)=23.36$, $P=0.001^*$	10.30 (3.28)
PANSS_distress (s.d.)	—	9.68 (2.60)	11.0 (3.77)	10.3 (3.28)	$F(1,84.4)=8.13$, $P=0.005^*$	12.94 (5.32)
PANSS_disorganization (s.d.)	—	12.4 (4.59)	13.5 (5.95)	12.9 (5.32)	$F(1,65.3)=2.33$, $P=0.13$	13.04 (5.41)
GAF (s.d.)	—	57.35 (14.8)	46.16 (11.93)	51.93 (14.58)	$F(1,218)=39.2$, $P<0.001^*$	65.98 (15.28)
BDI-II (s.d.)	—	11.008 (11.354)	9.718 (10.957)	10.3805 (11.074)	$F(1,171)=0.287$, $P=0.593$	15.96 (13.65)
SANS_blunted affect	—	0.510 (0.765)	1.07 (1.17)	0.793 (1.03)	$F(1,252)=23.99$, $P<0.001^*$	0.509 (0.85)
SANS_alogia	—	0.528 (0.797)	1.15 (1.24)	0.840 (1.08)	$F(1,249)=26.39$, $P<0.001^*$	0.519 (0.872)
SANS_avolition	—	0.250 (0.514)	0.988 (1.22)	0.618 (1.01)	$F(1,195)=45.28$, $P<0.001^*$	0.317 (0.70)
SANS_anhedonia	—	1.88 (1.79)	2.76 (1.88)	2.32 (1.89)	$F(1,291)=17.05$, $P<0.001^*$	1.133 (1.612)
SANS_asociality	—	1.09 (1.08)	1.33 (1.21)	1.29 (1.16)	$F(1,288)=8.91$, $P=0.003^*$	0.669 (0.971)
Independent validation sample						
Number, <i>n</i>	159	94	137	231		132
Age (years), mean (s.d.)	25.0 (5.33)	26.04 (5.47)	25.8 (6.12)	26.07 (5.71)	$F(2,231)=2.13$ $P=0.121$	26.24 (5.74)
Sex (<i>n</i> for male (female))	70 (89)	46 (48)	75 (62)	123 (110)	$F(2,233)=0.965$ $P=0.382$	123 (53.25%)
Ethnicity (%)					$F(2,286)=0.663$ $P=0.516$	
White	84.83	87.82	77.82	82.33		82.31
Black	2.04	—	4.04	2.04		3.04
Asian	3.43	0.71	5.53	3.43		3.40
Caribbean	0.1	0.1	0.0	1.44		—
Other	9.60	11.37	12.61	10.76		11.25
PANSS_positive (s.d.)	—	10.28 (2.95)	15.19 (4.03)	12.11 (4.14)	$F(1,116)=8107$, $P<0.001^*$	11.29 (5.48)
PANSS_negative (s.d.)	—	16.23 (7.19)	14.47 (6.71)	15.58 (7.05)	$F(1,158)=2.93$, $P=0.089$	13.38 (6.083)
PANSS_excitement (s.d.)	—	12.4 (3.96)	13 (4.26)	12.62 (4.08)	$F(1,140)=0.9368$, $P=0.335^*$	10.96 (3.64)
PANSS_distress (s.d.)	—	18.28 (6.27)	17.22 (5.48)	17.89 (5.99)	$F(1,165)=1.15$, $P=0.221$	13.78 (5.38)
PANSS_disorganization (s.d.)	—	13.59 (3.61)	16.97 (6.18)	14.85 (4.99)	$F(1,100)=17.85$, $P<0.001^*$	13.59 (5.42)
GAF (s.d.)	—	52.22 (12.38)	45.5 (16.36)	47.66 (15.48)	$F(1,90.6)=5.93$, $P<0.002^*$	62.62 (16.25)
BDI-II (s.d.)	—	24.99 (14.76)	16.413 (15.09)	20.249 (15.521)	$F(1,230)=14.36$, $P<0.001^*$	14.49 (13.02)
SANS_blunted affect	—	0.481 (0.672)	1.109 (1.34)	0.867 (1.16)	$F(1,65.5)=6.754$, $P=0.012$	0.917 (1.68)
SANS_alogia	—	0.593 (0.888)	1.27 (1.42)	1.01 (1.28)	$F(1,68)=6.050$, $P=0.016$	0.695 (1.34)
SANS_avolition	—	0.324 (0.7)	0.890 (1.27)	0.671 (1.11)	$F(1,67.1)=5.759$, $P=0.019$	0.404 (0.81)
SANS_anhedonia	—	2.59 (1.97)	2.23 (1.89)	2.37 (1.91)	$F(1,53.7)=0.573$, $P=0.452$	0.443 (0.83)
SANS_asociality	—	1.16 (1.24)	1.35 (1.32)	1.28 (1.28)	$F(1,57.8)=0.364$, $P=0.548$	1.057 (1.23)

ANOVA, analysis of variance; PANSS, positive and negative symptom scale; GAF, global assessment of functioning; BDI-II, Beck Depression Inventory-II; SANS (scale for the assessment of negative symptoms). * $P=0.01$.

Table 2 | Classification results

	Balanced accuracy (%)	Sensitivity (%)	Specificity (%)	Follow-up ROP+ROD prediction (%)
Training sample baseline	72.43	74.40	74.45	83.25
Independent validation sample baseline	60.27	61.20	58.65	56

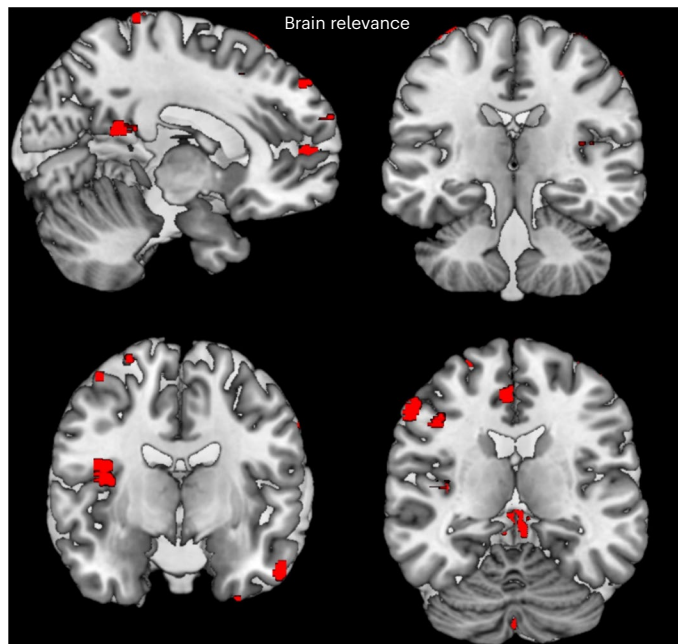


Fig. 1 | Voxels with the red color contribute more to the classification decision. The relevance of the voxels calculated from the LRP are overlaid and rendered in a GM template using the MRIcron toolbox.

Classification results and localization

A repeated nested pooled cross-validation classifier of brain contrast texture maps achieved a balanced accuracy of 72.43%, a sensitivity of 74.40% and a specificity of 74.45% for discrimination of patients (pooled ROD and ROP) from the HC participants (Table 2). A total of 231 age- and sex-matched participants was included in the independent validation sample, with 139 participants classified correctly as ROD + ROP (balanced accuracy 60.27%, sensitivity 61.20% and specificity 58.65%). In the independent validation follow-up sample, 137 participants were included, where 77 participants were classified correctly and 60 of them provided all the clinical variables.

Implementation of the layer-wise relevance propagation (LRP) algorithm for multilayer neural networks, as described in Bach et al.³⁸, showed that voxels with the highest contribution to the difference between patients and HC participants (that is, those with the highest relevance for the classification decision) were located in the cerebellum, anterior and posterior midline areas, frontotemporal areas and right insula (Fig. 1 and Supplementary Tables 5 and 6). In addition, we investigated cortical biomarkers for the general psychopathology in gray matter (GM), WM and CSF. The regions of interest were extracted using the AAL-VOIs atlas (<https://doi.org/10.1006/nimg.2001.0978>). We used the Johns Hopkins University (JHU) WM tractography atlas³⁹ to identify the WM tractography identified by the LRP.

Brain relevance association with clinical outcome profile

The association of the brain relevance with clinical symptoms and outcome profiles was investigated in the validation sample. Clinical

variables at the baseline were available for 131 patients from the validation sample (out of 139 classified correctly) and 71 patients of the follow-up sample. We support identifying patterns in the brain relevance heatmaps that were associated with the severity of clinical symptoms and outcome profiles. We applied the affinity propagation (AP) algorithm⁴⁰ to cluster the brain relevance heatmaps of the validation sample blindly^{14,37}, resulting in eight clusters. Inside each cluster, we calculated (1) the SANS factor scores (for the avolition, asociality, anhedonia, alogia and blunted affect subgroups), GAF and BDI total scores as well as the PANSS factor scores (for the positive, negative, distress, excitement and disorganization subgroups) at baseline (Fig. 2) and (2) the change scores of these variables at the one-year follow-up time (Fig. 3). The resulting clusters are presented in Figs. 2 and 3. We summarize the findings here on the basis of the symptom sum scores, as described above, and the outcome profile (details in Supplementary Table 7). Cluster 1 (number of subjects (*N*), *N* = 9 at baseline and *N* = 6 at follow-up) presented high BDI and PANSS_negative values, an improvement in PANSS_negative over time and a deterioration in PANSS_disorganization and the SANS scores related to alogia and blunted affect over time (Supplementary Fig. 3). In contrast to Cluster 1, Cluster 3 (*N* = 21 at baseline and *N* = 12 at follow-up) presented low BDI values and improvement in PANSS_disorganization as well as a high deterioration of the SANS_asociality and SANS_anhedonia scores and the BDI after 1 year. We speculate that the brain relevance in Cluster 3 has prognostic power for depression. Cluster 2 (*N* = 8 at baseline and *N* = 5 at follow-up) presented a high value for the GAF and low values for the PANSS_negative and SANS factor scores, with a marked improvement in the BDI and SANS_blunted affect and a deterioration in GAF, SANS_anhedonia and SANS_avolition over time. In contrast to Cluster 2, Cluster 4 (*N* = 12 at baseline and *N* = 6 at follow-up) presented low values in total GAF and high values in the SANS factor scores, with marked improvement in GAF, SANS_asociality and SANS_alogia over time and marked deterioration of the PANSS scores over time. We speculate that the negative symptoms in the baseline were a secondary effect of the primary positive symptoms. Cluster 6 (*N* = 31 at baseline and *N* = 17 at follow-up) and Cluster 7 (*N* = 6 at baseline and *N* = 5 at follow-up) captured older and younger ages, respectively. Cluster 7 presents a high improvement in the SANS_avolition and SANS_anhedonia scores. Finally, Cluster 5 (*N* = 22 at baseline and *N* = 9 at follow-up) presented low a SANS_blunted affect score and deterioration in the PANSS scores over time, and Cluster 8 (*N* = 22 at baseline and *N* = 11 at follow-up) showed a strong improvement in the PANSS_positive scores over time.

In clusters with a sufficient number of participants (Clusters 3, 5, 6 and 8), we repeated regression analyses to predict clinical outcomes at baseline and changes in the scores between baseline and follow-up, using the mean relevance of the brain heatmap as a predictor (in Table 1, the mean and standard deviation scores of the ROP and ROD participants in the follow-up are presented). In Cluster 3, the mean relevance map predicted an improvement of the PANSS score factors (positive, disorganization, distress and excitement) and the SANS_anhedonia score with negative coefficients (Supplementary Table 8.1). In Cluster 8, the mean relevance map predicted an improvement of the PANSS_positive score factor with a positive coefficient (Supplementary Table 8.2). In addition, we tested the prediction values of the relevance of the voxels that contributed most to the classification decision. There were differential associations between voxel relevance and symptoms in the different clusters, especially for Clusters 3, 5, 6 and 8. In Cluster 3, the voxels in specific brain areas that are associated with low depressive symptoms (that is, fusiform, temporal lobe and insula from the left hemisphere; see Supplementary Table 7) predicted PANSS scores (for the excitement, disorganization and positive subgroups) and an improvement of functionality. In Cluster 5, the brain relevance predicts the improvement of functionality from the left temporal lobe and in the right-hemisphere calcarine, lingual gyrus and precuneus regions. In Cluster 6, the brain relevance associated with high age predicts the

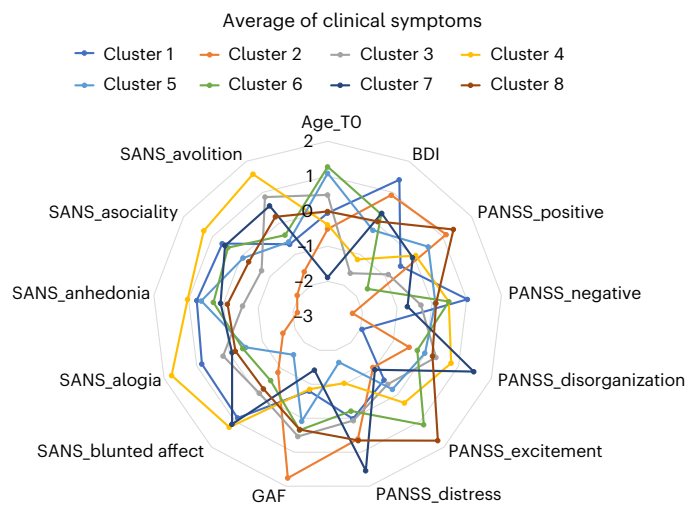


Fig. 2 | The average values of the clinical variables at baseline calculated in each cluster for the participants belonging to the independent validation sample. Different colors represent the different clusters.

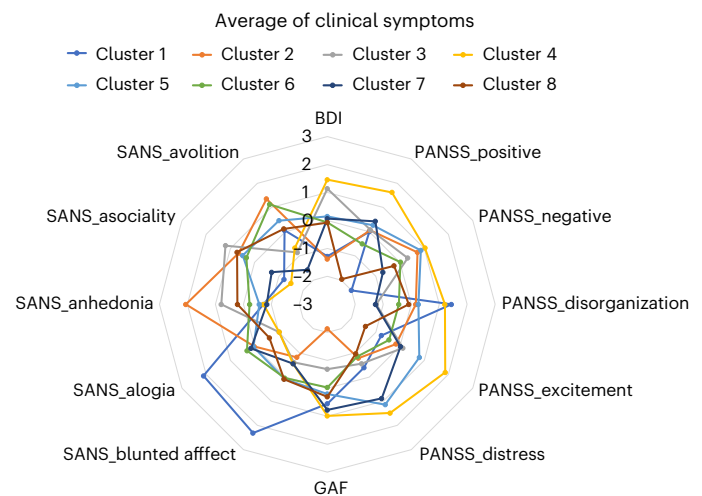


Fig. 3 | The average values of the difference in symptoms in one year after the baseline calculated in each cluster. Different colors represent the different clusters.

functionality at baseline from the left occipital and temporal lobes and the fusiform gyrus. In Cluster 8, the brain relevance predicts the improvement of functionality from the left temporal lobe.

Furthermore, we calculated the volumes of the regions with predicted value to check if the contribution of the voxels to the classification decision was driven by the volumetric changes. In all clusters with high numbers of participants (that is, Clusters 3, 5, 6 and 8), the brain volume in the detected regions was uncorrelated with the classification relevance of the respective regions, with the exception of the precentral and frontal lobe regions that are not contributing to the statistical models (Fig. 4).

Discussion

In this study, across patients with ROD or ROP we identified eight clusters of brain heatmap patterns that showed differential associations with symptom clusters, functionality and age; that is, clinical heterogeneity in transdiagnostic psychopathology was reflected in homogeneous clusters of contrast texture brain changes. The brain regions contributing most to the classifier decision overlap with those implicated in psychotic and depressive disorders in studies that have assessed changes in the GM and WM of patients³⁷. This fact and the proof-of-concept results of this study support our notion that neuroimaging biomarkers are useful for the prediction of transdiagnostic outcome profiles. We used the contrast feature map and the LRP method suggested by Bach et al.³⁸ to train and explain a classifier for identification of the transdiagnostic psychopathology. The model showed good accuracy in classifying a combined group of ROD and ROP patients against HC participants in the training sample, which decreased in the external validation sample. However, the innovative finding of this study is the creation of homogeneous clusters that are associated with clinical severity and outcome profiles. The current analysis provides an example of how the proposed method can be applied to link the transdiagnostic symptoms (rather than categorical diagnoses) to brain structure. Our ultimate vision pertains to multi-dimensional prediction, investigating neuroanatomical markers of longitudinal changes in symptoms and functionality across diagnoses.

Changes in the hippocampus, insula and lateral prefrontal cortex have been observed in patients with ROD (for example, a decrease in volume in medial temporal regions)²⁷ and ROP (for example, a volume deficit in the anterior hippocampus)⁴¹, which emphasizes the implication of brain alterations in the early stages of both illnesses. Using our

method, GM analysis in transdiagnostic psychopathology showed brain changes in regions that have been reported in other studies, for example, the cerebellum²⁰, the frontal lobe⁴², temporal regions⁴³, the cingulum bundle⁴⁴ and the precuneus⁴⁵, rectus⁴⁶, insula⁴⁷, Heschl's gyrus⁴⁸, lingual⁴⁹ and putamen left⁵⁰ regions. The explained classification results using the contrast feature maps indicate the posterior cingulate cortex, gyrus rectus, third ventricle (consistent with our previous results in patients with schizophrenia and major depressive disorder¹⁴) and insular (consistent with previous study for clinical high risk who transitioned to psychosis and first-episode psychosis^{37,51}) as key regions for the transdiagnostic classification of psychopathology (see Table 3 for further details).

In people with mental-health problems, there is mounting evidence of decreased WM function that results in problems of synchronization and connectivity⁵², with fractional anisotropy (FA) evaluated via diffusion tensor imaging being the most extensively used measure. Throughout the phases and progression of psychosis, neural alterations, particularly changes in WM connectivity, could be seen⁵³. Consistent findings of significantly lower FA in the left genu of the corpus callosum, the left anterior corona radiata and the right superior longitudinal fasciculus have been reported in ROP patients relative to HC participants^{33,54,55}. Functional dysconnectivity in the dorsal anterior cingulate cortex have been revealed in major depression⁵⁶. Expansion of the third ventricle has been reported in the recent onset of psychosis⁵⁷. Other studies have reported abnormalities in WM integrity as measured by FA in major depressive disorders³⁴. The above findings are in line with our results regarding the WM changes in the frontal gyrus, gyrus rectus, dorsal anterior and corpus callosum. The third and fourth ventricles have been reported as regions that contribute to the identification of the transdiagnostic psychopathology using the contrast texture map.

The brain relevance calculated using the LRP algorithm resulted in clusters that showed differential associations with age, positive, negative and depressive symptom clusters as well as functionality. Previous studies that performed sMRI-based prediction models determined social functioning in the patients with ROD²⁷. In some clusters, mean relevance values and/or the relevance of voxels contributing significantly to the classification decision accurately predicted the PANSS and SANS sum scores, functionality and change in functionality over time. Significant contributing voxels were located in regions that have previously been associated with these scores. In previous studies on schizophrenia and ROD, the FA of the temporal part and the temporo-lingual

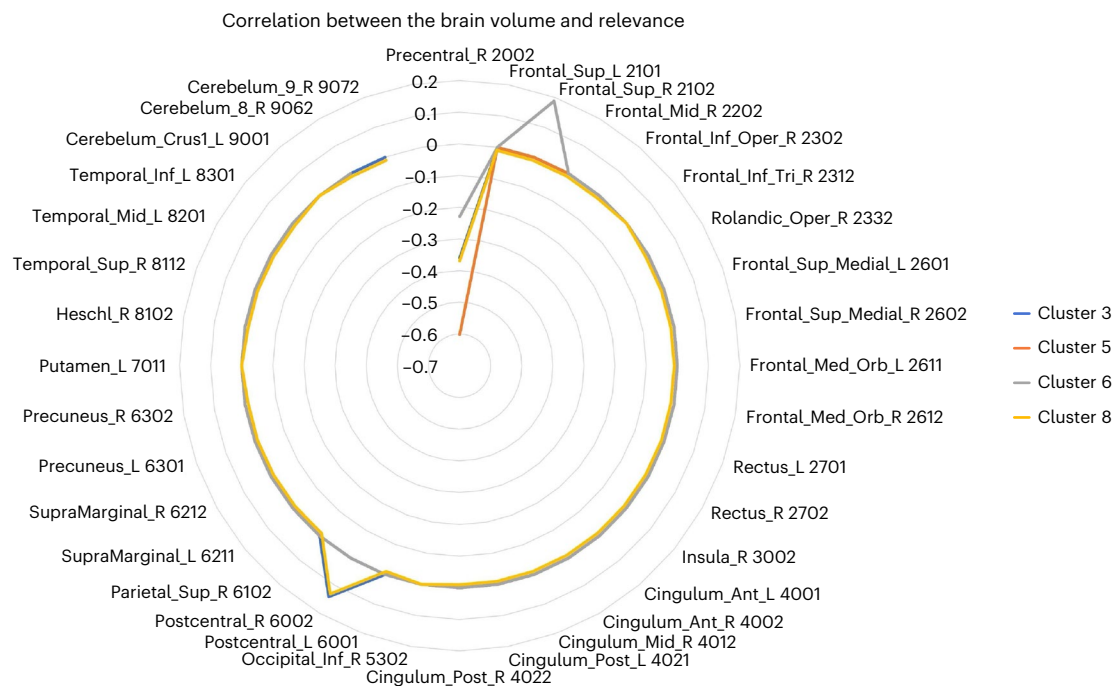


Fig. 4 | The correlation between the brain volume and the relevance of the voxels that contribute more to the classification decision in clusters that contain more than 20 participants. Different colors represent the different

clusters. Brain regions labels extracted by the AAL-VOIs atlas. L, left; R, right; Sup, superior; Mid, middle; Inf, inferior; Oper, operculum; Tri, triangularis; Med, medial; Orb, orbital; Ant, anterior; Post, posterior; Crus1, Crus1 lobule.

GM was positively correlated with the GAF score, consistent with the findings in Cluster 6, which includes 17 ROP and 14 ROD patients at baseline^{27,58}. A reduction of GM in the fusiform gyrus and temporal lobe were associated with PANSS scores in schizophrenia⁵⁹, as we observed in Cluster 3, which includes mainly ROP patients. An improvement in functionality has also been associated with a increased GM volume in temporal lobes at baseline in a previous study⁶⁰. Our proposed method yielded additional, new findings, which should be validated further in different datasets—for example, the right-hemisphere calcarine, lingual and precuneus regions for the prediction of functionality at follow-up, the fusiform gyrus and insula for the prediction of PANSS scores at baseline, and the occipital lobe and fusiform gyrus for the prediction of GAF scores at baseline.

The innovation of this study is the development and validation of individualized transdiagnostic models that give insights regarding the relation of brain changes to clinical symptoms and functional outcome. The identified clusters were derived using sMRI of recent samples and may eventually lead to the development of MRI-based predictive and decision-making tools. Given that the heterogeneity of mental disorders has hindered the search for biomarkers so far^{5,26,27}, we created homogeneous clusters based on brain relevance patterns. We found groups with different brain alteration profiles, which corresponded to different transdiagnostic clinical profiles while also showing distinct association profiles between clinical symptom clusters and anatomy. Larger studies are warranted to investigate the optimal number of stable clusters; for the purposes of this study, it is important to show that MRI can potentially be used to form a clinical categorization into more homogeneous groups than those offered by categorical classification systems^{3,61,62}. Distinct psychopathological profiles may be established to help distinguish patients with different syndromes, independent of other diagnostic considerations, and relate these to symptoms and clinical outcomes. Models that emphasize the role of changeable transdiagnostic illness mechanisms can help to further these efforts, promoting the generalizability of evidence-based treatments to routine care settings by accommodating comorbidities.

The extraction of texture feature maps from unsegmented brain images provides insight into the inter-relationships of different modality voxels. To our knowledge, non-segmented images have never been used to detect the examined disorders due to the lack of any interaction between GM, WM and CSF as an indicator of diagnosis, suggesting a novel biomarker^{14,37}. The main advantages of the proposed method are the interpretability of the results and the use of non-segmented images, which eliminate segmentation errors. It considers that the changes in contrast reveal microstructure changes when calculated in small three-dimensional cubes, and the image intensities reveal inter-relationships between GM, WM and CSF. The radiomic texture feature maps were extracted in MNI space (see Methods) from a registered masked T_1 -weighted image. The radius of the cubes and the topology of the neural network, for example, should be investigated further. For the replication of results, we used the same cube size and preprocessing methods as in previous papers^{14,37}. Future studies should investigate a range of textural qualities, as each one reflects a different component of brain diversity. Finally, diverse methodological techniques should be examined further to acquire a better understanding of neurobiological variations across disorders and to make the results useful for targeted interventions and treatment alternatives. Future studies in larger samples should investigate whether or not several neuroimaging modalities may be combined and used for more accurate prediction, as well as an investigation of whether or not the regions involved represent functional networks.

Limitations

Currently, psychiatric diagnosis is not based on underlying brain structures or clear biological etiology. Patients whose symptoms may be caused by different biological processes may receive the same diagnosis, and patients whose symptoms may be caused by the same biological process may receive different diagnoses. This practice may adversely affect the development of outcome predictions and explain the limit in transferability of the results. Results must be interpreted with a clear understanding of any limitations, such as the cross-sectional

and longitudinal nature of our data, which is important in interpretation because there may be dynamic and changing symptom cluster profiles that are not captured. The extraction of texture features is impacted by the variation in MRI intensity standardization. Another methodological drawback of the proposed method lies in the present lack of consensus regarding the applied image normalization method within the process for the extraction of texture features. It is important to note that, although the validation sample was recruited within the same study as the discovery sample, not all patients in the validation sample were recruited from the same sites as the discovery patients. Thus, environmental factors may have affected the accuracy prediction of our models. Because of this limitation, we emphasize that the present results represent a proof of concept, and that further validation in larger and geographically diverse cohorts is needed. Future studies should investigate thoroughly the stability and generalizability of the model as well as test the potential application of the model to clinical practice. Further investigation is needed on whether or not several neuroimaging modalities may be combined and used for more accurate prediction, as well as whether or not the regions involved represent functional networks.

Conclusions

The present analysis confirms that, at a symptom cluster level, the psychopathological profile of ROP and ROD follows similar patterns that can be clustered in homogeneous groups associated with specific brain alterations. We identified grouped transdiagnostic neuroanatomical signatures, which were strongly related with homogeneous clinical severity and outcome profiles. The present findings support a hypothesis of common psychopathological profiles of ROP and ROD and recognition of the outcome profile that could inform the development of novel targeted and repurposed therapies that target specific symptom clusters rather than specific diagnoses.

Methods

Before being included in the study, written informed consent was obtained from all participants and guardians (in the case of underage participants (defined as those under 18 years of age at all sites)). The trial was registered with the German Clinical Trials Register (DRKS00005042), and local research ethics boards in each location gave their approval (Ludwig-Maximilian University Munich (ethics ID: 351-13), University of Basel (ethics ID: M12/99), University of Cologne (ethics ID: 13-236), University of Turku (ethics ID: 99/1810/2013), University of Bari (ethics ID: 4754), University of Milan (ethics ID: N.PROT.0001885|PIGEN/02), University of Udine (ethics ID: 67172), University of Birmingham (ethics ID: 14/WM/0019), University of Münster (ethics ID: 2016-398-b-S) and University of Düsseldorf (ethics ID: 5957A)). Participants could claim transport costs. The STROBE (strengthening the reporting of observational studies in epidemiology) checklist is included in Supplementary Fig. 1.

Study design

Data were collected in the context of the PRONIA study, a seven-center study funded by the Seventh Framework Programme of the European Union aimed at optimizing candidate biomarkers for the prediction of mental disorders in early states of the disease. Details of the PRONIA study sites, recruitment protocol and quality-control procedures have been described in detail in previous publications^{26,27}. We refer the reader to Koutsouleris et al.²⁷ for a detailed description of the study design and methodology.

Participants

Data were collected following the standardized recruitment and assessment protocol from the PRONIA (PRONIA) study across seven European sites for the training sample (Ludwig-Maximilian University Munich (209 consenting participants), University of Basel (98 consenting

participants), University of Cologne (143 consenting participants), University of Turku (93 consenting participants), University of Milan (42 consenting participants), University of Birmingham (84 consenting participants) and University of Udine (92 consenting participants) (Supplementary Table 1) and across ten European sites for the validation sample (Ludwig-Maximilian University Munich (158 consenting participants), University of Basel (41 consenting participants), University of Cologne (65 consenting participants), University of Turku (67 consenting participants), University of Milan (47 consenting participants), University of Birmingham (40 consenting participants), University of Udine (51 consenting participants), University of Düsseldorf (31 consenting participants), University of Münster (45 consenting participants) and University of Bari (26 consenting participants)). Race and ethnicity were reported to the interviewer based on the individual's identification. In both samples across diagnoses, white people comprised more than 77% of the total. In brief, participants were recruited through outpatient and inpatient services. The study participants analysed in the present study were recruited following a standardized recruitment and ascertainment protocol. The observational study protocol involved follow-up examinations every three months after the index ascertainment. Regular inter-rater reliability tests were performed to ensure reliability of the GAF scales and PANSS scores across the study sites. PRONIA investigators were repeatedly trained by one of the authors of the GAF and PANSS scales, A. Auther, who was independently testing the PRONIA consortium using four written transcripts of interviews from Zucker Hillside Hospital, USA. Thirty-six PRONIA raters performed an intraclass correlation (ICC) analysis to measure the between-rater agreement on the target measures. For each reliability test the raters had to generate six functioning scores: the current, lowest and highest in the past year for the social and role functioning domain (GAF:S/GAF:R) and for the PANSS scores. The ICC analysis results for the GAF scores are presented in Supplementary Table 2. Cicchetti⁶³ gives the following guidelines for interpretation of the kappa or ICC inter-rater agreement measures: <0.40, poor; 0.40–0.59, fair; 0.60–0.74, good; 0.75–1.00, excellent. The psychopathology of each participant was assessed by trained clinicians. Calibration of the PANSS[ICC = 0.79] scores was performed by implementing inter-rater reliability tests across the study sites. The inter-rater reliability test for the SANS scores is not available⁶⁴. The following were the general inclusion criteria: (1) in the age range of 15–40 years old, (2) sufficient language skills for participation and (3) the ability to provide informed consent. Current or previous head trauma with loss of consciousness (>5 min), current or past known neurological or somatic disorders that potentially affect the structure or functioning of the brain, current or past alcohol dependence, polysubstance dependence within the past six months and any medical indication against MRI were all general exclusion criteria. The following were all additional exclusion criteria for the HC participants: any current or previous Axis I or II disorder as defined in the *Diagnostic and Statistical Manual of Mental Disorders* 4th Edition (DSM-IV); a positive family history for affective or non-affective psychoses in first-degree relatives; and psychotropic medication intake more than five times per year and in the month before inclusion in the study.

ROP participants had to meet the same requirements as participants for a first episode of affective or non-affective psychotic episode, including the transition criteria established by Yung et al.⁶⁵ and the structured clinical interview for DSM-IV-TR (SCID)⁶⁶. They also had to be less than three months removed from the start of their initial antipsychotic medication treatment. The following specific ROP exclusion criteria comprised psychosis onset occurring more than 24 months ago and antipsychotic use lasting longer than 90 days (cumulatively in the past 24 months) with a daily dose rate at or above the 'first-episode psychosis' range minimum dosage of the German Society for Psychiatry, Psychotherapy and Nervous Diseases (DGPPN) S3 guidelines, equivalent to 5 mg olanzapine.

Patients with ROD were required to meet the SCID-established serious depression criteria that had been met within the previous three months. The specific ROD exclusion criteria were any DSM-IV-TR major depressive episode that preceded the current or recent episode, a duration of more than 24 months for the the current episode or the use of antipsychotic medication for more than 30 days, as described above.

Training and validation samples

The total training sample of 435 age- and sex-matched participants included 116 with ROD, 122 with ROP (minimally treated first-episode psychosis), and 197 HC, recruited during the first part of the recruitment period. All models were validated in the same participants at follow-up after 12 months, as well as in an independent validation sample (137 ROD, 162 ROP and 178 HC) recruited during the second part of the recruitment period. The follow-up sample consisted of 100 patients with ROP and 103 patients with ROD. Matching the validation sample to the discovery sample for sex and age resulted in a final independent validation sample of 94 ROD, 137 ROP and 159 HC participants, which was used in this study (details in Supplementary Fig. 1). The follow-up sample consisted of 81 patients with ROP and 48 patients with ROD. To evaluate the minimum sample size needed to have an experiment that is sensitive enough to detect the specified hypothetical effect size, we ran a power analysis.

Assessments

Data for the present analysis included the following: demographic and clinical data information (age, sex and medication exposure), Beck Depression Inventory-II (BDI-II)⁶⁷, positive and negative symptom scale (PANSS)⁶⁸, global assessment of functioning (GAF; global functioning-role scales)⁶⁹ and the scale for the assessment of negative symptoms (SANS). We calculated five subscores for PANSS (positive, negative, disorganized, excitement and distress) based on the factor analysis by Wallwork and co-workers⁷⁰. Several factor-analytic studies have suggested that a five-factor model better captures the PANSS structure in schizophrenia samples⁷⁰. To verify that the consensus factor structure proposed by Wallwork et al. (2012)⁷⁰, was suitable for our mixed sample of ROP and ROD patients, we performed confirmatory factor analysis. The results confirmed that the five PANSS factors had an acceptable fit to the data, which was reflected in a comparative fit index greater than 0.9 for all five factors (Supplementary Table 3). For the GAF, we used ratings based exclusively on the level of functioning; symptoms were not considered. The SANS factor scores provide a more detailed and differentiated picture of clinical symptoms, some of which are not included in the negative-factor PANSS, while others may have varying courses. To verify that the consensus factor structure proposed by Strauss et al. (2018)⁷¹ was suitable for our mixed sample of ROP and ROD patients, we performed confirmatory factor analysis. The results confirmed that the hierarchical model of the SANS factors had an acceptable fit to the data, which was reflected in a comparative fit index of greater than 0.9 for all four factors and shows significance for the anhedonia subgroup (Supplementary Table 4).

All participants underwent sMRI (at a field strength of 1.5 T (14 ROP, 7 ROD and 13 HC for the training sample and 12 ROP, 9 ROD and 9 HC for the validation sample), with the rest at a field strength of 3 T)^{25,26}. For the current analysis, we used T_1 -weighted sMRI images of the ROD and ROP participants. In keeping with real-world scanner heterogeneity and as part of the larger PRONIA goals, the PRONIA sites were required (1) to acquire isotropic or nearly isotropic voxel sizes with a resolution of at least 1 mm, (2) to set the field-of-view parameters accordingly to ensure full three-dimensional coverage of the brain, including all parts of the cerebellum and (3) to define the relaxation time and echo time as well as other ionizing radiation parameters. At every site all images were equally distributed across the field strength, visually inspected, automatically defaced and anonymized using an in-house FreeSurfer-based script (<https://surfer.nmr.mgh.harvard.edu>) before

the data was centralized. Supplementary Table 1 lists the scanner and parameter details of the structural magnetic resonance sequences used to examine the PRONIA sample participants. See Supplementary Section 1.1 and the previous PRONIA report²⁶ for full MRI harmonization and data acquisition parameters.

Analysis

Preprocessing. All images were visually inspected, automatically defaced and anonymized using a FreeSurfer-based script before data centralization. Subsequently, we used the open-source CAT12 toolbox (version r1155; CAT12), an extension of SPM12 (Statistical Parametric Mapping: The Analysis of Functional Brain Images, SPM12), to segment the images into GM, WM and CSF maps and to high-dimensionally register the segmented images to the stereotactic space of the Montreal Neurological Institute coordinates (MNI-152 space). The CAT12 toolbox was used with processing steps that consisted of spatial filtering, segmentation, segmentation estimation, a local adaptive segmentation step (which adjusts the images for WM homogeneities and varying GM intensities) and a high-dimensional Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL) registration of the image to an MNI template in the Ixi database (IXI). The manual of the CAT12 toolbox details the processing steps applied to the structural images. The quality of the GM volume maps was checked using the established quality-assurance framework of the CAT12 toolbox (details in Supplementary Section 1.1)³⁷. The images were inspected and removed if artifacts were identified during the quality-assurance framework of CAT12. The non-segmented magnetic resonance images, that is, $wp0^*$ images, were used for further analysis. We used histogram equalization to adjust the contrast of a grayscale image. Most pixel values of the original image present low contrast, and are in the middle of the intensity range. The `histeq` function in MATLAB R2020 produces an output image with pixel values evenly distributed throughout the grayscale range. The texture feature maps were then extracted from the transformed $wp0^*$ images (details in Supplementary Section 1.2 and Supplementary Fig. 2).

Feature extraction

We extracted texture feature maps from non-segmented images using the two-dimensional gray-level co-occurrence matrix (GLCM) calculated in each cube ($N = 3,294$ feature maps in total). Based on a previous study by Korda and colleagues³⁷, we extracted the texture features 'entropy', 'sum of entropy', 'difference of entropy', 'energy', 'contrast' and 'homogeneity'. We focused on the analysis in GLCM-contrast, which gives a low weight to elements with similar gray-level values and a high weight to elements with dissimilar gray levels, indicating large differences between neighboring voxels³⁷. In our previous independent study, contrast feature maps were used for the identification of first-episode psychosis patients resulting in the highest accuracy in previous research³⁷. Contrast captures high differences in gray levels in a small neighborhood. The highest difference is observed in the cortical boundary and at the borders between segments. Smaller differences observed in brain regions revealed the shape abnormalities that are not always related to volumetric changes. For this reason, we selected contrast feature maps for further investigation. The contrast of the gray-level pairs reflects intracortical myelin as has been investigated for patients with schizophrenia in low-sensory and motor areas⁷² (see Supplementary Fig. 3 for the whole preprocessing).

To test the transdiagnostically discrimination power of the ROP and ROD patients against the HC participants, the registered texture feature maps on the MNI space were fed into a 20×20 nested cross-validation deep-learning scheme for group classification (Supplementary Fig. 4). We focus the description of results on the contrast feature map, which gives a low weight to elements with similar gray-level values and a high weight to elements with dissimilar gray levels (that is, indicating large differences between neighboring voxels⁷³).

Classification framework

As we were interested in transdiagnostic psychopathology, the ROP and ROD patients were pooled together for individualized brain texture analysis. We fed the registered contrast texture feature maps in a multilayer neural network in MATLAB R2020b to train and cross-validate a model to discriminate transdiagnostically the ROP and ROD patients from the HC participants (see Supplementary Fig. 4 for further details). Feature selection (two-sample *t*-test) in the inner cycle was cross-validated by selecting a number of features appropriate to the dimension of the database, namely, the top 549 ranked features that best discriminated the two classes⁷⁴. A total of 549 features was selected using cross-validated feature selection to reduce dimensionality in the data. The number of features was chosen to balance the number of the features and the number of participants (curse of dimensionality).

Visualization and evaluation of heatmaps

For localization of the identified brain changes between pooled ROP and ROD versus HC participants, we calculated the relevance of the voxels in each class using the LRP algorithm for multilayer neural networks, as described by Bach and co-workers³⁸. The output of the algorithm is an individual heatmap that represents common changes in brain structure for general psychopathology (see Appendix A in the Supplementary Information for more details).

The final images were visualized using the MRICron toolbox (MRICron; v1.020190902 for Windows). Visualizations of the classification results on the holdout dataset are presented in Fig. 1. The regions of interest were extracted using the AAL-VOIs atlas (AAL-VOIs). GM, WM and CSF were investigated for cortical biomarkers in psychosis. We used the JHU WM tractography atlas³⁹ to identify the WM tractography identified by the LRP algorithm.

Associations of clusters with clinical variables

We implemented a clustering algorithm to demonstrate shared brain texture patterns across the participants. Our intention was to display the heatmap of each correctly classified participant in the independent validation sample. The independent validation sample was tested in the winner model of the training sample. The relevance of the voxels produced by the LRP algorithm was clustered in groups; subsequently, we associated the clusters with symptom severity and outcome profile⁷⁵. The AP algorithm⁴⁰ (Supplementary Section 1.4) was selected to cluster the participant's relevance heatmap for identifying distinct patterns of the brain changes in the transdiagnostic psychopathology.

In addition, we performed simple and multiple regression analysis to investigate and model clinical symptoms and outcome profiles, inside the clusters, from (1) the mean brain relevance (heatmap) and (2) the contributed voxels, in the independent validation baseline and the follow-up holdout dataset.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data used/analysed in the current study are available from the authors upon reasonable request submitted to the PRONIA Consortium. Data can only be shared on request due to restrictions laid out in the study's participant consent forms. Data will be made available from N.K. upon reasonable request submitted to the PRONIA Consortium (Nikolaos.Koutsouleris@med.uni-muenchen.de) after a data-sharing agreement with the PRONIA Consortium has been signed.

Code availability

The codes used in the current study are available from N.K. upon reasonable request submitted to the PRONIA Consortium (Nikolaos.Koutsouleris@med.uni-muenchen.de). The mathematical equations

for calculation of the contrast texture feature and the LRP algorithm are provided in the Supplementary Information.

References

- Menchetti, M., Murri, M. B., Bertakis, K., Bortolotti, B. & Berardi, D. Recognition and treatment of depression in primary care: effect of patients' presentation and frequency of consultation. *J. Psychosom. Res.* **66**, 335–341 (2009).
- van Os, J. et al. Schizophrenia and the environment: within-person analyses may be required to yield evidence of unconfounded and causal association—the example of cannabis and psychosis. *Schizophr. Bull.* **47**, 594–603 (2021).
- Fusar-Poli, P. et al. Transdiagnostic psychiatry: a systematic review. *World Psychiatry* **18**, 192–207 (2019).
- Newman, D. L., Moffitt, T. E., Caspi, A. & Silva, P. A. Comorbid mental disorders: implications for treatment and sample selection. *J. Abnorm. Psychol.* **107**, 305–311 (1998).
- Schultze-Lutter, F., Schmidt, S. J. & Theodoridou, A. Psychopathology—a precision tool in need of re-sharpening. *Front. Psychiatry* 10.3389/fpsyt.2018.00446 (2018).
- Lalousis, P. A. et al. Transdiagnostic structural neuroimaging features in depression and psychosis: a systematic review. *Neuroimage Clin.* **38**, 103388 (2023).
- Correll, C. U. & Schooler, N. R. Negative symptoms in schizophrenia: a review and clinical guide for recognition, assessment, and treatment. *Neuropsychiatr. Dis. Treat.* **16**, 519–534 (2020).
- Sönmez, N., Romm, K. L., Andreassen, O. A., Melle, I. & Røssberg, J. I. Depressive symptoms in first episode psychosis: a one-year follow-up study. *BMC Psychiatry* **13**, 106 (2013).
- Häfner, H. et al. Schizophrenia and depression: challenging the paradigm of two separate diseases—a controlled study of schizophrenia, depression and healthy controls. *Schizophr. Res.* **77**, 11–24 (2005).
- Iqbal, Z., Birchwood, M., Chadwick, P. & Trower, P. Cognitive approach to depression and suicidal thinking in psychosis. 2. Testing the validity of a social ranking model. *Br. J. Psychiatry* **177**, 522–528 (2000).
- Upthegrove, R. Depression in schizophrenia and early psychosis: implications for assessment and treatment. *Adv. Psychiatr. Treat.* **15**, 372–379 (2009).
- Upthegrove, R., Marwaha, S. & Birchwood, M. Depression and schizophrenia: cause, consequence, or trans-diagnostic issue? *Schizophr. Bull.* **43**, 240–244 (2017).
- Goodkind, M. et al. Identification of a common neurobiological substrate for mental illness. *JAMA Psychiatry* **72**, 305–315 (2015).
- Korda, A. I. et al. Identification of voxel-based texture abnormalities as new biomarkers for schizophrenia and major depressive patients using layer-wise relevance propagation on deep learning decisions. *Psychiatry Res. Neuroimaging* **313**, 111303 (2021).
- Kirschner, M. et al. Ventral striatal dysfunction and symptom expression in individuals with schizotypal personality traits and early psychosis. *Schizophr. Bull.* **44**, 147–157 (2018).
- Depping, M. S., Schmitgen, M. M., Kubera, K. M. & Wolf, R. C. Cerebellar contributions to major depression. *Front. Psychiatry* **9**, 634 (2018).
- Borgwardt, S. J. et al. Structural brain abnormalities in individuals with an at-risk mental state who later develop psychosis. *Br. J. Psychiatry* **191**, s69–s75 (2007).
- Chopra, S. et al. Differentiating the effect of antipsychotic medication and illness on brain volume reductions in first-episode psychosis: a longitudinal, randomised, triple-blind, placebo-controlled MRI study. *Neuropsychopharmacology* **46**, 1494–1501 (2021).

19. Busatto, G. F. Structural and functional neuroimaging studies in major depressive disorder with psychotic features: a critical review. *Schizophrenia. Bull.* **39**, 776–786 (2013).
20. Meurs, M. et al. The associations of depression and hypertension with brain volumes: independent or interactive? *Neuroimage Clin.* **8**, 79–86 (2015).
21. Pantelis, C. et al. Neuroanatomical abnormalities before and after onset of psychosis: a cross-sectional and longitudinal MRI comparison. *Lancet* **361**, 281–288 (2003).
22. Crespo-Facorro, B. et al. Caudate nucleus volume and its clinical and cognitive correlations in first episode schizophrenia. *Schizophr. Res.* **91**, 87–96 (2007).
23. Fusar-Poli, P., Radua, J., McGuire, P. & Borgwardt, S. Neuroanatomical maps of psychosis onset: voxel-wise meta-analysis of antipsychotic-naïve VBM studies. *Schizophr. Bull.* **38**, 1297–1307 (2012).
24. Koutsouleris, N. et al. Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. *Arch. Gen. Psychiatry* **66**, 700–712 (2009).
25. Lalouis, P. A., et al. Neurobiologically based stratification of recent onset depression and psychosis: identification of two distinct transdiagnostic phenotypes. *Biol. Psychiatry* 10.1016/j.biopsych.2022.03.021 (2022).
26. Lalouis, P. A. et al. Heterogeneity and classification of recent onset psychosis and depression: a multimodal machine learning approach. *Schizophr. Bull.* **47**, 1130–1140 (2021).
27. Koutsouleris, N. 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* **75**, 1156–1172 (2018).
28. Koutsouleris, N. et al. Multimodal machine learning workflows for prediction of psychosis in patients with clinical high-risk syndromes and recent-onset depression. *JAMA Psychiatry* **78**, 195–209 (2021).
29. Arnone, D., McIntosh, A. M., Tan, G. M. Y. & Ebmeier, K. P. Meta-analysis of magnetic resonance imaging studies of the corpus callosum in schizophrenia. *Schizophr. Res.* **101**, 124–132 (2008).
30. Vita, A., De Peri, L., Deste, G. & Sacchetti, E. Progressive loss of cortical gray matter in schizophrenia: a meta-analysis and meta-regression of longitudinal MRI studies. *Transl. Psychiatry* **2**, e190–e190 (2012).
31. Lai, C.-H. & Wu, Y.-T. Frontal-insula gray matter deficits in first-episode medication-naïve patients with major depressive disorder. *J. Affect. Disord.* **160**, 74–79 (2014).
32. Sørensen, N. V., Orlovskaa-Waast, S., Jeppesen, R., Christensen, R. H. & Benros, M. E. Neuroimmunological investigations of cerebrospinal fluid in patients with recent onset depression – a study protocol. *BMC Psychiatry* **22**, 35 (2022).
33. Walterfang, M. et al. White matter volume changes in people who develop psychosis. *Br. J. Psychiatry* **193**, 210–215 (2008).
34. Shen, X. et al. Subcortical volume and white matter integrity abnormalities in major depressive disorder: findings from UK Biobank imaging data. *Sci. Rep.* **7**, 5547 (2017).
35. Giné-Servén, E. et al. Routine cerebrospinal fluid parameters as biomarkers in first-episode psychosis: a prospective observational study. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **112**, 110424 (2022).
36. Falkai, P., Schmitt, A. & Andreasen, N. Forty years of structural brain imaging in mental disorders: is it clinically useful or not? *Dialogues Clin. Neurosci.* **20**, 179–186 (2018).
37. Korda, Alexandral. et al. Identification of texture MRI brain abnormalities on first-episode psychosis and clinical high-risk subjects using explainable artificial intelligence. *Transl. Psychiatry* **12**, 481 (2022).
38. Bach, S. et al. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS ONE* **10**, e0130140 (2015).
39. Hua, K. et al. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *Neuroimage* **39**, 336–347 (2008).
40. Frey Brendan, J. & Dueck, D. Clustering by passing messages between data points. *Science* **315**, 972–976 (2007).
41. McHugo, M. et al. Hippocampal volume in early psychosis: a 2-year longitudinal study. *Transl. Psychiatry* **10**, 306 (2020).
42. Reig, S. et al. Progression of brain volume changes in adolescent-onset psychosis. *Schizophr. Bull.* **35**, 233–243 (2009).
43. Veijola, J. et al. Longitudinal changes in total brain volume in schizophrenia: relation to symptom severity, cognition and antipsychotic medication. *PLoS ONE* **9**, e101689 (2014).
44. Whitford, T. J. et al. Localized abnormalities in the cingulum bundle in patients with schizophrenia: a diffusion tensor tractography study. *Neuroimage Clin.* **5**, 93–99 (2014).
45. Li, G., Rossbach, K., Zhang, A., Liu, P. & Zhang, K. Resting-state functional changes in the precuneus within first-episode drug-naïve patients with MDD. *Neuropsychiatr. Dis. Treat.* **14**, 1991–1998 (2018).
46. Zheng, R., Zhang, Y., Yang, Z., Han, S. & Cheng, J. Reduced brain gray matter volume in patients with first-episode major depressive disorder: a quantitative meta-analysis. *Front. Psychiatry* **12**, 671348 (2021).
47. Hatton, S. N. et al. Correlating anterior insula gray matter volume changes in young people with clinical and neurocognitive outcomes: an MRI study. *BMC Psychiatry* **12**, 45 (2012).
48. Dickey, C. C. et al. Smaller left Heschl's gyrus volume in patients with schizotypal personality disorder. *Am. J. Psychiatry* **159**, 1521–1527 (2002).
49. Zhang, T., Koutsouleris, N., Meisenzahl, E. & Davatzikos, C. Heterogeneity of structural brain changes in subtypes of schizophrenia revealed using magnetic resonance imaging pattern analysis. *Schizophr. Bull.* **41**, 74–84 (2015).
50. Lu, Y. et al. The volumetric and shape changes of the putamen and thalamus in first episode, untreated major depressive disorder. *Neuroimage Clin.* **11**, 658–666 (2016).
51. Li, X.-B. et al. Altered resting-state functional connectivity of the insula in individuals with clinical high-risk and patients with first-episode schizophrenia. *Psychiatry Res.* **282**, 112608 (2019).
52. van Velzen, L. S. et al. White matter disturbances in major depressive disorder: a coordinated analysis across 20 international cohorts in the ENIGMA MDD working group. *Mol. Psychiatry* **25**, 1511–1525 (2020).
53. Carletti, F. et al. Alterations in white matter evident before the onset of psychosis. *Schizophr. Bull.* **38**, 1170–1179 (2012).
54. Surbeck, W. et al. Anatomical integrity within the inferior fronto-occipital fasciculus and semantic processing deficits in schizophrenia spectrum disorders. *Schizophr. Res.* **218**, 267–275 (2020).
55. Barth, C. et al. Exploring white matter microstructure and the impact of antipsychotics in adolescent-onset psychosis. *PLoS ONE* **15**, e0233684 (2020).
56. Ho, T. C. et al. Inflexible functional connectivity of the dorsal anterior cingulate cortex in adolescent major depressive disorder. *Neuropsychopharmacology* **42**, 2434–2445 (2017).
57. Chung, Y. et al. Prodromal symptom severity predicts accelerated gray matter reduction and third ventricle expansion among clinically high-risk youth developing psychotic disorders. *Complex Psychiatry* **1**, 13–22 (2015).
58. Yamada, S. et al. Cognitive and functional deficits are associated with white matter abnormalities in two independent cohorts of patients with schizophrenia. *Eur. Arch. Psychiatry Clin. Neurosci.* **272**, 957–969 (2022).

59. Mennigen, E. et al. Positive and general psychopathology associated with specific gray matter reductions in inferior temporal regions in patients with schizophrenia. *Schizophr. Res.* **208**, 242–249 (2019).
60. Kambeitz-Ilankovic, L. et al. Multivariate pattern analysis of brain structure predicts functional outcome after auditory-based cognitive training interventions. *npj Schizophr.* **7**, 40 (2021).
61. Haslam, N., McGrath, M. J., Viechtbauer, W. & Kuppens, P. Dimensions over categories: a meta-analysis of taxometric research. *Psychol. Med.* **50**, 1418–1432 (2020).
62. Markon, K. E., Chmielewski, M. & Miller, C. J. “The reliability and validity of discrete and continuous measures of psychopathology: a quantitative review”: Correction to Markon et al. (2011). *Psychol. Bull.* **137**, 1093–1093 (2011).
63. Cicchetti, D. V. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol. Assess.* **6**, 284–290 (1994).
64. Oeztuerk, O. F. 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.* **272**, 403–413 (2022).
65. Yung, A. R. et al. Prediction of psychosis: a step towards indicated prevention of schizophrenia. *Br. J. Psychiatry* **172**, 14–20 (1998).
66. Regier, D. A., Kuhl, E. A. & Kupfer, D. J. The DSM-5: classification and criteria changes. *World Psychiatry* **12**, 92–98 (2013).
67. Beck, A. T., Steer, R. A. & Brown, G. *Beck Depression Inventory-II* (American Psychological Association, 1996).
68. Kay, S. R., Fiszbein, A. & Opler, L. A. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* **13**, 261–276 (1987).
69. Andreasen, N. C. The Scale for the Assessment of Negative Symptoms (SANS): conceptual and theoretical foundations. *Br. J. Psychiatry* **155**, 49–52 (1989).
70. Wallwork, R. S., Fortgang, R., Hashimoto, R., Weinberger, D. R. & Dickinson, D. Searching for a consensus five-factor model of the Positive and Negative Syndrome Scale for schizophrenia. *Schizophr. Res.* **137**, 246–250 (2012).
71. Strauss, G. P. et al. The latent structure of negative symptoms in schizophrenia. *JAMA Psychiatry* **75**, 1271–1279 (2018).
72. Jørgensen, K. N. et al. Increased MRI-based cortical grey/white-matter contrast in sensory and motor regions in schizophrenia and bipolar disorder. *Psychol. Med.* **46**, 1971–1985 (2016).
73. Brynolfsson, P. et al. Haralick texture features from apparent diffusion coefficient (ADC) MRI images depend on imaging and pre-processing parameters. *Sci. Rep.* **7**, 4041 (2017).
74. Zanetti, M. V. et al. Neuroanatomical pattern classification in a population-based sample of first-episode schizophrenia. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **43**, 116–125 (2013).
75. Stelzer, J., Lacosse, E., Bause, J., Scheffler, K. & Lohmann, G. Brainlance: visualizing group level MRI data at one glance. *Front. Neurosci.* 10.3389/fnins.2019.00972 (2019).

Acknowledgements

A.I.K., C.A. and S.B. are grateful for computational support and infrastructure provided by the OMICS at the University of Lübeck (Germany). Preliminary results were presented at ECNP conference 2022 by A.I.K., who also achieved the Excellence Award. PRONIA is a Collaboration Project funded by the European Union under the Seventh Framework Programme and grant agreement 602152. The sponsors had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; and decision to submit the manuscript for publication.

Author contributions

A.I.K. conceived and designed the study, feature extraction and analysis, and contributed to interpretation of the results and writing the manuscript. C.A. conceived and designed the study and contributed to interpretation of the results and writing the paper. A.R. and L.H. were responsible for preprocessing of the brain MRI. A.S., U.D., L.K.-I., D.B.D., J.K., J.W., S.R., S.J.W., R.K.R.S., C.P., F.S.-L., A.B., J.H., E.M., P.B., P.S., R.U., P.A.L., A.R.-R., C.D., R.L., S.B. and the PRONIA Consortium were responsible for data acquisition and support on writing the manuscript. N.K. had full access to all the data in the study and takes responsibility for the integrity of the data.

Funding

Open access funding provided by Universität zu Lübeck.

Competing interests

N.K. received grants from the European Union (EU) during the conduct of the study and has a patent (US20160192889A1) issued. S.R. received grants from the European Commission during the conduct of the study. A.R.-R. received grants from the EU during the conduct of the study. C.A. received non-financial support from Sunovion Pharmaceuticals, Inc. and H. Lundbeck A/S outside the submitted work. A.R.-R. received grants from The Zurich Program for Sustainable Development of Mental Health Services (for Zurich Early Recognition Program (ZInEP)) and support from a private donation. C.P. received grants from Australian National Health and the Medical Research Council during the conduct of the study and personal fees from H. Lundbeck A/S and Australia Pty Ltd outside the submitted work. R.U. received personal fees from Sunovion Pharmaceuticals, Inc. outside the submitted work. E.M. has a patent (US20160192889A1) licensed. No other disclosures were reported.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44220-023-00168-z>.

Correspondence and requests for materials should be addressed to Alexandra I. Korda.

Peer review information *Nature Mental Health* thanks Michael Breakspear and the other, anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

Alexandra I. Korda^{1,24}✉, Christina Andreou^{1,24}, Anne Ruef², Lisa Hahn², André Schmidt³, Udo Dannlowski⁴, Lana Kambeitz-Ilankovic⁵, Dominic B. Dwyer², Joseph Kambeitz⁵, Julian Wenzel⁵, Stephan Ruhrmann⁵, Stephen J. Wood^{6,7,8}, Raimo K. R. Salokangas⁹, Christos Pantelis¹⁰, Frauke Schultze-Lutter^{11,12,13}, Eva Meisenzahl¹¹, Paolo Brambilla^{14,15}, Pierluigi Selvaggi¹⁶, Alessandro Bertolino¹⁷, Jarmo Hietala⁹, Rachel Upthegrove^{18,19}, Paris Alexandros Lalouis¹⁹, Anita Riecher-Rössler²⁰, Christos Davatzikos²¹, Rebekka Lencer^{1,4}, Nikolaos Koutsouleris^{2,22,23}, Stefan Borgwardt^{1,3} & the PRONIA Consortium*

¹Department of Psychiatry and Psychotherapy and Center for Brain, Behaviour and Metabolism, University of Luebeck, Luebeck, Germany. ²Department of Psychiatry and Psychotherapy, Ludwig-Maximilian University Munich, Munich, Germany. ³Department of Psychiatry, Psychiatric University Hospital, University of Basel, Basel, Switzerland. ⁴Institute for Translational Psychiatry and Otto Creutzfeldt Center for Cognitive and Behavioral Neuroscience, University of Muenster, Muenster, Germany. ⁵Department of Psychiatry and Psychotherapy, Faculty of Medicine and University Hospital, University of Cologne, Cologne, Germany. ⁶School of Psychology, University of Birmingham, Birmingham, UK. ⁷Orygen, Melbourne, Victoria, Australia. ⁸Centre for Youth Mental Health, University of Melbourne, Melbourne, Australia. ⁹Department of Psychiatry, University of Turku, Turku, Finland. ¹⁰Melbourne Neuropsychiatry Centre, Department of Psychiatry, University of Melbourne & Melbourne Health, Melbourne, Victoria, Australia. ¹¹Department of Psychiatry and Psychotherapy, Medical Faculty, Heinrich-Heine University, Duesseldorf, Germany. ¹²Department of Psychology, Faculty of Psychology, Airlangga University, Surabaya, Indonesia. ¹³University Hospital of Child and Adolescent Psychiatry and Psychotherapy, University of Bern, Bern, Switzerland. ¹⁴Department of Neurosciences and Mental Health, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy. ¹⁵Department of Pathophysiology and Transplantation, University of Milan, Milan, Italy. ¹⁶Department of Translational Biomedicine and Neuroscience (DiBrain), University of Bari Aldo Moro, Bari, Italy. ¹⁷Department of Basic Medical Science, Neuroscience, and Sense Organs, University of Bari Aldo Moro, Bari, Italy. ¹⁸Birmingham Early Intervention Service, Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK. ¹⁹Institute for Mental Health and Centre for Human Brain Health, School of Psychology, University of Birmingham, Birmingham, UK. ²⁰Medical Faculty, University of Basel, Basel, Switzerland. ²¹Department of Radiology, University of Pennsylvania School of Medicine, Philadelphia, PA, USA. ²²Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. ²³Max Planck Institute of Psychiatry, Munich, Germany. ²⁴These authors contributed equally: Alexandra I. Korda, Christina Andreou. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: alexandra.korda@uksh.de

the PRONIA Consortium

Alexandra I. Korda^{1,24}, Christina Andreou^{1,24}, Anne Ruef², Lisa Hahn², André Schmidt³, Udo Dannlowski⁴, Lana Kambeitz-Ilankovic⁵, Dominic B. Dwyer², Joseph Kambeitz⁵, Julian Wenzel⁵, Stephan Ruhrmann⁵, Stephen J. Wood^{6,7,8}, Raimo K. R. Salokangas⁹, Christos Pantelis¹⁰, Frauke Schultze-Lutter^{11,12,13}, Eva Meisenzahl¹¹, Paolo Brambilla^{14,15}, Pierluigi Selvaggi¹⁶, Alessandro Bertolino¹⁷, Jarmo Hietala⁹, Rachel Upthegrove^{18,19}, Paris Alexandros Lalouis¹⁹, Anita Riecher-Rössler²⁰, Christos Davatzikos²¹, Rebekka Lencer^{1,4}, Nikolaos Koutsouleris^{2,22,23} & Stefan Borgwardt^{1,3}

A full list of members and their affiliations appears in the Supplementary Information.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☐ ☒ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The data used/analyzed in the current study will be available from the authors upon reasonable request submitted to the PRONIA Consortium. Data can only be shared on request due to restrictions laid out in the study's participant consent forms. Data will be made available after a data-sharing agreement with PRONIA Consortium has been signed. Estimated timeframe is 3 months. Contact: Nikolaos Koutsouleris, email: Nikolaos.Koutsouleris@med.uni-muenchen.de

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	The term sex reported in this study
Population characteristics	The population description can be found in the manuscript Table 1 and 2.
Recruitment	Data were collected following the standardized recruitment and assessment protocol from the Personalised pROgNostic tools for early psychosis management (PRONIA; https://www.pronia.eu) study across 7 European sites: Ludwig-Maximilian University Munich (157 participants consent), University of Basel (79 participants consent), University of Cologne (116 participants consent), University of Turku (72 participants consent), University of Milan (32 participants consent), University of Birmingham (71 participants consent) (Supplementary Table S1). The ethnicity was identified by the examiner at the baseline. In brief, participants were recruited through outpatient and inpatient services. The study participants analysed in the present study were recruited following a standardized recruitment and ascertainment protocol. The observational study protocol involved follow-up examinations every three months after the index ascertainment.
Ethics oversight	Prior to being included in the study, all participants, and, in case of underage participants (defined as those under the age of 18 at all sites), and their guardians gave written informed consent. The trial was registered with the German Clinical Trials Register (DRKS00005042), and local research ethics boards in each location gave their approval. Written informed consent was obtained from all participants. Participants could claim transport costs.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The power analysis on the training sample showed that 94 ROP+ROD and 78 HC is the minimum sample size which result in power greater than 0.9.
Data exclusions	No Data were excluded
Replication	We applied a 10 times repeated 20x20 nested cross validation approach in the training sample which was validates in the independent validation sample.
Randomization	We pooled the data of ROP and ROD together and fed into the 10 times repeated 20x20 nested cross validation framework.
Blinding	The ROP and ROD participants from all the sites pooled together and fed into the explainable AI approach.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input type="checkbox"/>	<input checked="" type="checkbox"/> MRI-based neuroimaging

Magnetic resonance imaging

Experimental design

Design type	structural brain MRI
Design specifications	N/A
Behavioral performance measures	N/A

Acquisition

Imaging type(s)	structural
Field strength	1.5T and 3T
Sequence & imaging parameters	Table S1 in the Supplement file
Area of acquisition	whole brain scan
Diffusion MRI	<input type="checkbox"/> Used <input checked="" type="checkbox"/> Not used

Preprocessing

Preprocessing software	The preprocessing steps were executed in the PRONIA server at LMU by Prof. Nikolaos Koutsouleris
Normalization	Normalization of the brain images was performed.
Normalization template	a high-dimensional DARTEL registration of the image to a MNI-template generated from the MRI data of 555 healthy controls in the IXI database (http://www.braindevelopment.org). The registered GM images were multiplied with the Jacobian determinants obtained during registration to produce GM volume (GMV) maps. Images were smoothed with 10 mm before entering the subsequent analysis steps. (8) The Quality Assurance framework of CAT12 was used to check the quality of the GMV maps.
Noise and artifact removal	Processing steps applied to the structural images, consisting of (1) the 1st denoising step based on Spatially Adaptive Non-Local Means (SANLM) filtering 2;(2) an Adaptive Maximum A Posteriori (AMAP) segmentation technique, which models local variations of intensity distributions as slowly varying spatial functions and thus achieves a homogeneous segmentation across cortical and subcortical structures3; (3) the 2nd denoising step using Markov Random Field approach which incorporates spatial prior information of adjacent voxels into the segmentation estimation generated by AMAP3; (4) a Local Adaptive Segmentation (LAS) step, which adjusts the images for white matter (WM) inhomogeneities and varying gray matter (GM) intensities caused by the differing iron content in e.g. cortical and subcortical structures. The LAS step is carried out before the final AMAP segmentation; (5) a partial volume segmentation algorithm that is capable of modeling tissues with intensities between GM and WM, as well as GM and cerebrospinal fluid (CSF) and is applied to the AMAP-generated tissue segments
Volume censoring	The Quality Assurance framework of CAT12 was used to check the quality of the GMV maps.

Statistical modeling & inference

Model type and settings	Explainable AI, deep learning
Effect(s) tested	n/a
Specify type of analysis:	<input type="checkbox"/> Whole brain <input type="checkbox"/> ROI-based <input type="checkbox"/> Both
Statistic type for inference (See Eklund et al. 2016)	voxel-wise
Correction	FDR corrected p-value

Models & analysis

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Functional and/or effective connectivity
<input checked="" type="checkbox"/>	<input type="checkbox"/> Graph analysis
<input type="checkbox"/>	<input checked="" type="checkbox"/> Multivariate modeling or predictive analysis

Multivariate modeling and predictive analysis	We applied a 10 times repeated 20x20 nested cross validation approach in the training sample which was validated in the independent validation sample. 549 features were selected using a cross-validated t-test.
---	---