Psychosis brain subtypes validated in first-episode cohorts and related to illness remission: Results from the PHENOM consortium

Dominic B Dwyer
Ludwig-Maximilian University
Ganesh B Chand
Washington University School of Medicine in St. Louis
Aristeidis Sotiras
Washington University School of Medicine in St. Louis
et al.

Follow this and additional works at: https://digitalcommons.wustl.edu/oa_4

Part of the Medicine and Health Sciences Commons

Please let us know how this document benefits you.

Recommended Citation
Dwyer, Dominic B; Chand, Ganesh B; Sotiras, Aristeidis; and et al., "Psychosis brain subtypes validated in first-episode cohorts and related to illness remission: Results from the PHENOM consortium." Molecular Psychiatry. 28, 5. 2008 - 2017. (2023).
https://digitalcommons.wustl.edu/oa_4/3078

This Open Access Publication is brought to you for free and open access by the Open Access Publications at Digital Commons@Becker. It has been accepted for inclusion in 2020-Current year OA Pubs by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
Using machine learning, we recently decomposed the neuroanatomical heterogeneity of established schizophrenia to discover two volumetric subgroups—a ‘lower brain volume’ subgroup (SG1) and an ‘higher striatal volume’ subgroup (SG2) with otherwise normal brain structure. In this study, we investigated whether the MRI signatures of these subgroups were also already present at the time of the first-episode of psychosis (FEP) and whether they were related to clinical presentation and clinical remission over 1-, 3-, and 5-years. We included 572 FEP and 424 healthy controls (HC) from 4 sites (Sao Paulo, Santander, London, Melbourne) of the PHENOM consortium. Our prior MRI subgrouping models (671 participants; USA, Germany, and China) were applied to both FEP and HC. Participants were assigned into 1 of 4 categories: subgroup 1 (SG1), subgroup 2 (SG2), no subgroup membership (‘None’), and mixed SG1 + SG2 subgroups (‘Mixed’). Voxel-wise analyses characterized SG1 and SG2 subgroups. Supervised machine learning analyses characterized baseline and remission signatures related to SG1 and SG2 membership. The two dominant patterns of ‘lower brain volume’ in SG1 and ‘higher striatal volume’ (with otherwise normal neumorphology) in SG2 were identified already at the first episode of psychosis. SG1 had a significantly higher proportion of FEP (32%) vs. HC (19%) than SG2 (FEP, 21%; HC, 23%). Clinical multivariate signatures separated the SG1 and SG2 subgroups (balanced accuracy = 64%; p < 0.0001), with SG2 showing higher education but also greater positive psychosis symptoms at first presentation, and an association with symptom remission at 1-year, 5-year, and when timepoints were combined. Neuromorphological subtypes of schizophrenia are already evident at illness onset, separated by distinct clinical presentations, and differentially associated with subsequent remission. These results suggest that the subgroups may be underlying risk phenotypes that could be targeted in future treatment trials and are critical to consider when interpreting neuroimaging literature.
INTRODUCTION

Biological heterogeneity underlying psychosis has prevented the identification and implementation of precision medicine approaches in the clinical care of this disorder. A long history of research into the relationship between brain morphology and diagnoses [1], illness stages [2], or symptoms [3] has used top-down approaches, but the lack of clinical translation [4–6] has motivated calls for bottom-up approaches to identify neuromorphological subgroups directly from brain scans [7–12]. Brain clustering is one approach that aims to find naturally occurring subgroupings of individuals defined by similar brain patterns using brain volume [7] or cortical thickness measures [11]. Such studies can identify subgroups that explain underlying psychosis heterogeneity in the neuroimaging field and lead to more precise treatment approaches if related to clinical outcomes, such as symptom remission.

To date, studies have focused on determining neuroimaging subgroups within samples of patients with psychosis in advanced illness stages [7–11, 13], reporting two [7, 8, 11, 13, 14], three [10, 15, 16], and six [17] subgroup solutions. Such studies generally find separations between groups based largely on the severity of the brain differences in the patient population related to controls —e.g., with one subgroup demonstrating widespread reductions in volume [7] or thickness [10], and the other showing fewer or no reductions [7]. Evidence also suggests that some of the elements crucial to the subgrouping may not be specific to psychosis, as similar elements are found across diagnoses [11], in depression risk samples [18], and to some degree in samples of healthy controls [10, 14, 17]. These findings point to the presence of a brain-risk phenotype that is normatively present in the healthy population but is enriched in samples of individuals with psychosis. However, a major limitation in this research has been the limited internal or external validation of the subgroups in independent samples. Furthermore, commonly used clustering algorithms used for the subgroupings are often confounded by disease-unrelated factors that influence brain variability across individuals, such as age and sex [19].

The recent PHENOM consortium aimed to address these limitations by applying a semi-supervised machine learning technique (heterogeneity through discriminative analysis; HYDRA [19]) to cluster brain MRI volume in a large sample of schizophrenia patients with internal validation (i.e., cross, split-half, and leave-site-out validation) [7]. In this work, we found a stable and highly reproducible two-subgroup solution. One subgroup of patients (SG1) demonstrated widespread smaller cortical volume [7], as also found in previous work [8, 13]. For the first time, we also found a subgroup consisting of approximately one-third of patients that showed no cortical reductions compared to healthy controls but presented with larger striatal volumes (SG2) [7]. While these morphological alterations were not associated with baseline symptoms, individuals in SG2 exhibited higher educational attainment.

In our recent follow-up work, the PHENOM subgroups were also found to be expressed in population samples both in young adults (16–23 yrs), where greater SG1 membership was found in youth with psychosis spectrum symptoms, and in adults (44–50 years) where SG1 was associated with lower cognitive performance and higher genetic risk for schizophrenia [20]. These findings suggest that SG1 is associated with normative psychosis risk that is enriched in clinical samples while SG2 is not and may even be protective [21]. A limitation of ours and other work thus far, however, is that we do not know how these subgroups are expressed at the time of a first-episode of psychosis (FEP) when there is a confirmed psychosis diagnosis but less influence of illness or treatment confounds. Determination of FEP subgroup membership is also critical because this is a clinical window when decisions may influence the symptom and functioning illness course (i.e., prognoses) [22].

Longitudinal studies examining outcomes in first-episode populations have found substantial heterogeneity, especially in trans-diagnostic samples consisting of primary diagnoses of schizophrenia, bipolar disorder, and depression with psychotic symptoms [23, 24], with consistent reports of a subgroup of individuals (~30%) with unremitting symptoms [25] in addition to a diversity of other episodic courses [24, 25]. Such unremitting courses have been previously associated with male gender, lower educational achievement, and schizophrenia or non-affective diagnoses [22], together with widespread reductions in brain volume, gyrification abnormalities [23, 25–27], and less efficient connectivity [28, 29]. In contrast, studies have found that patients who have experienced periods of remission also have larger striatal volumes, similar to what we have seen in SG2, and particularly in female individuals [28, 29].

Such top-down clinical findings suggest that our SG1 (‘lower volume’) subgroup would be less likely to achieve remission when compared to the SG2 (‘increased striatum’) subgroup, potentially via relationships with female gender, higher educational achievement, and a diagnosis of psychosis different from schizophrenia. Supporting the first hypothesis, our previous cross-sectional study found an association with illness duration in the SG1 subgroup [7] and other preliminary research found an association between relapses and a cortical thinning subgroup [10]. However, a clear limitation of existing studies is that they mostly used a cross-sectional design with a mix of patients with first-episode and chronic psychosis [8, 10]. Research is thus required that specifically investigates remission longitudinally, beginning from the very first episode of illness.

Here, we applied the pre-trained subgroup models from chronic schizophrenia to an international multi-site sample consisting of individuals with first-episode psychosis (FEP) and healthy controls. First, we determined the proportion of FEP individuals (relative to HC) that belonged to the SG1 and SG2 subgroups with the expectation that the proportion of SG1 would be higher in FEP. Second, we explored baseline clinical signatures of the subgroups with the hypothesis that SG2 would demonstrate higher educational attainment. Third, we investigated remission over 1-, 3-, and 5-year longitudinal periods with the hypothesis that subjects in SG2 would be more likely to experience illness remission.

METHODS

Participants

We previously [7] created a two-group structural MRI statistical model in a sample with established schizophrenia from USA (n = 96), Germany (n = 145), and China (n = 66) in addition to controls from the same sites (n = 364). For the current study, we included a PHENOM subsample of 572 first-episode psychosis (FEP) patients, collected from 4 sites: Sao Paulo (n = 128; from two independent studies), Santander (n = 186), London (n = 122), and Melbourne (n = 136) (Supplemental Methods). In addition, 424 healthy control (HC) participants without Axis-I diagnoses were included from the same sites (Supplemental Methods). Sample size determination was based on discriminative effect sizes from our previous publications for machine learning imaging analyses [7, 20] and an expected 10–15% difference between subgroup proportions in baseline FEP patients and longitudinal remission [7, 20, 28]. We intentionally included samples that employed diverse sample recruitment, inclusion/exclusion criteria, and imaging protocols to assess the ability of the previously identified subgroups to generalize to a wide variety of contexts (see Supplementary Materials). Of note, none of these sites were included in the original training sample used to build the subgrouping models [7]. Local ethics committees at each site approved the studies and informed consent was obtained for all participants. Images were pooled at the Center for Biomedical Image Computing and Analytics of the University of Pennsylvania, USA.

Sociodemographic and clinical assessments

Clinical measures included a harmonized set of basic demographic and clinical legacy data from each site collected as part of the PHENOM consortium. Demographic measures included age, sex, and proxies for functioning including: relationship status (single or not), employment status (no/yes), and highest education (1 = 12 yrs; 2 = 16 yrs; 3 = 18 yrs; 4 = tertiary education).
Baseline and longitudinal univariate tests of subtype differences

Baseline differences in demographic characteristics, education, symptoms, and variables were assessed across all 4 subgroups using the F-test (continuous) and Chi-square/Kruskall-Wallis (categorical) analyses. Remission was analyzed comparing subgroups for each time-point using chi-squared analyses. In addition, a composite measure quantifying whether each individual available had ever remitted during any follow-up period (i.e., to identify individuals who exhibit ongoing long-term clinical impairment) was created. Longitudinal symptom data, as measured by the PANSS and the GAF-Symptoms, were also available for a subset of the sites and were analyzed with ANOVA. Two-sided tests were used for all analyses.

Supervised machine learning analyses separating subgroups and prediction of clinical outcomes

In order to investigate baseline clinical signatures separating the four subgroup classes (SG1, SG2, SG1 + SG2 and None) a multi-group machine learning analysis was conducted employing a nested cross-validation design (25-fold outer loop; fivefold inner loop with 4 shuffled permutations) including all clinical variables outlined in Table 1. Within each training fold, the clinical data were scaled and imputed using a k-nearest neighbors approach (7 neighbors). The preprocessed training data were then forwarded to an L1-regularized SVM classifier (LIBLINEAR: $C = (2^p \gamma \in \{-6,-4,\ldots,4\})$) to determine separation boundaries for each pair-wise brain subgroup. These models were applied without modification to the inner loop test data for all hyperparameters. Optimized models were applied without modification to the outer-loop held-out test data to obtain final accuracies. Permutation testing was conducted (10,000 iterations; labels swapped) to obtain significance levels for balanced accuracy estimates. The cross-validation ratio [39] and sign-based consistency [40] of feature weight measures were used to determine variable importance and significance. The same clinical variables and parameters within the multigroup setting were also employed to predict remission at each 1-, 3-, and 5-year timepoints and to determine any incidence of remission across all follow-up periods (i.e., nested cross-validation, preprocessing, and L1-regularized SVM). Based on our hypotheses regarding remission, the binary labels representing SG1, SG2, or SG1 + SG2 subgroup membership were added to the clinical measures in addition to the continuous SG1 and SG2 membership scores (see Fig. 1; we included both binary and continuous scores in the context of L1-regularized variable selection to determine which was maximally predictive).

RESULTS

Baseline and follow-up sample characteristics

Compared to controls, FEP cases had higher rates of single marital status and unemployment, but did not differ in age, sex, or education (Table S3). Substantial site differences were found across most demographic and clinical measures, except sex, the presence of a diagnosis of psychosis not otherwise specified (NOS), and duration of untreated psychosis (Table S4) (Supplemental Methods). For example, Melbourne included patients who were 5-years younger and had higher psychosis symptom severity (PANSS Total) compared to other sites, while Sao Paulo contained more cases with a diagnosis of schizophrenia (PSYCLASS sample) and lower educational attainment (PSYCLASS/ESNA samples). Follow-up participants were significantly older and had a higher prevalence of schizophrenia diagnoses, but less symptom load overall, and more use of illicit substances (Table S5).

Validation of neuromorphological subgroups at the time of the first psychotic episode

Site effects were effectively corrected (Figs. S1 & S2). Following application of our previously trained HYDRA model [7] we
<table>
<thead>
<tr>
<th></th>
<th>N % of total FEP</th>
<th>None</th>
<th>SG1</th>
<th>SG2</th>
<th>Mixed</th>
<th>F/Chi2/KW</th>
<th>p value</th>
<th>Eta2/Phi</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (data available)</td>
<td>572</td>
<td>217 (38)</td>
<td>184 (32)</td>
<td>118 (21)</td>
<td>53 (9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sites</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>London (n; %)</td>
<td>55 (25.3)</td>
<td>32 (17.4)</td>
<td>24 (20.3)</td>
<td>11 (20.8)</td>
<td>3.87 (8)</td>
<td></td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Melbourne (n; %)</td>
<td>58 (26.7)</td>
<td>39 (21.2)</td>
<td>30 (25.4)</td>
<td>9 (17.0)</td>
<td>3.25 (8)</td>
<td></td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Santander (n; %)</td>
<td>58 (26.7)</td>
<td>54 (29.3)</td>
<td>52 (44.1)</td>
<td>22 (41.5)</td>
<td>13.28 (8)</td>
<td></td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>Sao Paolo (n; %)</td>
<td>46 (21.2)</td>
<td>59 (32.1)</td>
<td>12 (10.2)</td>
<td>11 (20.8)</td>
<td>20.32 (8)</td>
<td></td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>Age (mean;SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex (male, %)</td>
<td>131 (60.4)</td>
<td>127 (69.0)</td>
<td>74 (62.7)</td>
<td>36 (67.9)</td>
<td>3.68 (8)</td>
<td></td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Onset Age (yrs; mean (SD))</td>
<td>26.1 (7.3)</td>
<td>26.0 (6.6)</td>
<td>26.1 (7.3)</td>
<td>25.0 (6.1)</td>
<td>1.05 (3568)</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of Illness (years; mean (SD))</td>
<td>0.7 (1.1)</td>
<td>0.7 (1.1)</td>
<td>0.7 (1.1)</td>
<td>1.3 (1.7)</td>
<td>2.62 (3417)</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAF (mean (SD))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PANSS Positive (mean (SD))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PANSS Negative (mean (SD))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PANSS General (mean (SD))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PANSS Total (mean (SD))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatments and Drug Use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPZ Z-norm (Mean (SD))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Duration (mths; mean (SD))</td>
<td>3.0 (2.5)</td>
<td>3.3 (2.6)</td>
<td>2.9 (2.2)</td>
<td>4.1 (4.8)</td>
<td>1.07 (3226)</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Lithium (yes, %)</td>
<td>424</td>
<td>11 (7.2)</td>
<td>6 (4.3)</td>
<td>8 (8.9)</td>
<td>2 (4.8)</td>
<td>2.35 (8)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Other mood stabilizers (yes, %)</td>
<td>182</td>
<td>1 (1.8)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>2.26 (8)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Antidepressants (yes, %)</td>
<td>284</td>
<td>7 (7.6)</td>
<td>9 (9.0)</td>
<td>3 (5.0)</td>
<td>5 (15.6)</td>
<td>3.18 (8)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Antiepileptics (yes, %)</td>
<td>102</td>
<td>1 (2.8)</td>
<td>6 (12.8)</td>
<td>0 (0.0)</td>
<td>2 (18.2)</td>
<td>4.52 (8)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Antipsychotic class typical/ atypical (typical, %)</td>
<td>335</td>
<td>13 (11.2)</td>
<td>23 (21.7)</td>
<td>2 (2.7)</td>
<td>6 (15.4)</td>
<td>14.42 (8)</td>
<td>2.38e–03</td>
<td>0.16</td>
</tr>
<tr>
<td>Marijuana (yes, %)</td>
<td>314</td>
<td>32 (30.8)</td>
<td>39 (34.5)</td>
<td>29 (45.3)</td>
<td>14 (42.4)</td>
<td>4.31 (8)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Other illicit (yes, %)</td>
<td>329</td>
<td>37 (33.9)</td>
<td>38 (34.2)</td>
<td>27 (37.5)</td>
<td>9 (24.3)</td>
<td>1.94 (8)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Tobacco (yes, %)</td>
<td>181</td>
<td>26 (46.4)</td>
<td>33 (64.7)</td>
<td>27 (51.9)</td>
<td>12 (54.5)</td>
<td>3.74 (8)</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Results corrected for multiple comparisons at a false-discovery rate (FDR) < 0.05. ANOVA used for continuous variables, Chi2 for binary, and Kruskall–Wallis for education variable. PANSS positive and negative syndrome scale, GAF global assessment of functioning, CPZ chlorpromazine equivalent dose, MDD major depressive disorder, Psychosis NOS Psychosis not otherwise specified.
obtained the following subgroup split: FEP (n = 572): SG1, n = 184 (32%); SG2, n = 118 (21%); SG1 + SG2, n = 53 (9%); None, n = 217 (34%); HC (n = 424): SG1, n = 82 (19%); SG2, n = 96 (23%); SG1 + SG2, n = 19 (5%); None, n = 227 (54%) (Fig. 1, Table 1). The higher proportion of FEP in SG1 compared to the other subgroups was significant (X²(8) = 36.5, p = 5.96e−08). These findings thus confirm that the two neuromorphological subtypes we had identified in patients in the advanced illness stages are already present at illness onset.

When whole-brain voxel-wise maps of the SG1 subgroup were compared to those of the ‘None’ subgroup in FEP, the SG1 subgroup showed smaller widespread cortical volumes coupled with smaller volume in some parts of caudate (Fig. 2). In contrast, larger gray matter volumes in SG2 subgroup were restricted to subcortical structures including the striatum. White matter was also smaller in SG1, mainly in subcortical regions adjoining the striatum, whereas white matter was larger in SG2 (Fig. 2). Because healthy control participants were also included in the SG1 and SG2 subgroups, we tested whether these subgroups demonstrated similar patterns to the patients and our previous work in population samples. Voxel-wise analyses were thus repeated in healthy controls comparing each subgroup definition to the ‘None’ subgroup to visualize the separation indicated by the application of the subgroup models (Figs. S3 and S4). These analyses demonstrated an SG1 pattern that was restricted in spatial extent and magnitude when compared to that of the full sample, although larger striatal volumes were also found in the HC of in SG2 (Fig. S3) and the pattern was no different from that of FEP patients (Fig. S4). Results were maintained when controlling for antipsychotic dose and type (i.e., atypical vs. typical; Supplementary Materials ‘Investigation of Medication Effects’; Fig. S5) and when investigating volume differences within individual sites (Fig. S6). Supplementary analyses in a small sample of SG1 + SG2 individuals showed both decreased cortical volume and increased striatal volume (Fig. S7). Global and distributed volume reductions in SG1 were maintained when controlling for intracranial volume (Fig. S8).
Differences in clinical baseline characteristics between SG1 and SG2 subgroups in FEP and HC

Subgroups were not different in age or sex, but a higher proportion of SG1 cases came from the Sao Paulo sample (i.e., containing more individuals with a schizophrenia diagnoses and lower education) and a higher proportion of SG2 cases came from the Santander site (Table 1). Importantly, these results occurred in the context of our site correction. The SG1 and SG1 + SG2 subgroups also included more individuals with a diagnosis of schizophrenia (SG1, 35%; SG1 + SG2, 50.9%). Individuals in SG1 (both FEP and HC) were also more likely to have a lower educational attainment at uncorrected levels (Table S6). SG1 were also prescribed proportionately more typical antipsychotics (Table 1).

Prediction of subgroup membership from clinical signatures

In patients, multigroup machine learning analyses revealed that the highest accuracy of subgroup separation based on clinical variables (see below) was associated with the SG1 vs. SG2 comparison (Table 2). SG2 membership was separated from SG1 with a sensitivity of 66% (balanced accuracy (BAC) = 64.03%) by a pattern including: education, higher positive symptoms, unemployment, female sex, and higher CPZ dose (z-scored within each site to account for site differences) (Fig. 3). Longer duration of untreated psychosis, diagnoses (including comorbidity), general psychosis symptoms, and family history of psychosis significantly predicted SG1 membership. Other subgroup comparisons were non-significant (although a trend was noted for SG1 vs. SG1 + SG2; Table 2).

Illness remission across subgroups

Individuals in SG2 were more likely to have experienced at least one period of symptom remission across all follow-up timepoints (SG1, 57%; SG2, 78%; $X^2(8) = 9.8, p = 0.02$; Table S7). For individual follow-up points, higher 1-year remission in SG2 was found at uncorrected levels (SG1, 52.7%; SG2, 73.2%; $X^2(8) = 7.7, p = 0.05$; Table S7). To provide a complementary perspective on the binary subgroup labels we also examined the relationship between membership strength (as the continuous decision scores) and likelihood of remission. We found that SG2 membership scores were significantly higher in cases who demonstrated at least one period of remission across all timepoints ($t(352) = -2.97, p = 0.003$), while there were no differences for SG1.

Prediction of remission from clinical variables and subgroup membership

Supervised machine learning was used to predict remission across timepoints using the clinical variables in addition to the SG1, SG2, and SG1 + SG2 labels to determine if subgroup membership was associated with more positive outcomes. The 'None' subgroup label was not included as it was not included as a hypothesis and we did not find sufficient separation at baseline. At the 1-year follow-up point (Santander/Sao Paulo), the balanced accuracy (BAC) was 64.2% (sensitivity, 60%; specificity, 68.5%; Table 3). The pattern predicting higher likelihood of remission included membership in SG2, increased schizophreniform disorder diagnoses, increased single marital status, and reduced schizophrenia diagnoses and unemployment (Fig. 4). While subgroup membership was not among baseline variables that predicted remission at 3-years (Santander), membership of SG2 predicted remission status at 5-years (London/Sao Paulo; BAC = 59.4%; sensitivity, 51.6%; specificity, 67.2%), together with a diagnosis of psychosis NOS, and female sex, and reduced likelihood of a schizophrenia diagnosis and single marital status. At 5-years (London/Sao Paulo; BAC = 59.4%; sensitivity, 51.6%; specificity, 67.2%) the pattern included higher SG2 membership, increased psychosis NOS, and female sex in addition to less schizophrenia diagnoses and single marital status.

Supplementary analyses of combined timepoints, site, antipsychotics, schizophrenia diagnosis, drug use, and clustering algorithms

Combining remission information across timepoints increases sample size for machine learning predictions. Remission
prediction at any time across 1-, 3-, and 5-year periods was predicted by a pattern that included increased likelihood of SG2 membership and decreased likelihood of SG1 membership at uncorrected levels, using the sign-based consistency measure of variable significance (Fig. S9).

Site differences were controlled for in brain volume measures (Figs. S1 and S2) prior to obtaining the subgroup labels. However, clinical analyses were conducted without site control because of clinical differences in inclusion/exclusion criteria, study populations, and experimental designs (Supplementary Materials), which could meaningfully influence membership into subgroups and the clinical prediction of remission. Controlling clinical analyses for site in this case therefore can control for important clinical variance related to outcomes. However, we repeated the analyses while controlling for site to determine if relationships with brain subgroups were maintained. Even when we controlled for site, SG2 membership remained predictive of remission at across all timepoints (Fig. S9) and at 5-years (Fig. S10), in addition to 1-year at uncorrected levels (Fig. S10). We also controlled for antipsychotic effects (dose and type) and the main findings were largely unchanged (Supplementary Materials ‘Investigation of Medication Effects’; Figs. S11 and S12).

In our analyses, a diagnosis of schizophrenia was a consistent feature associated with lower likelihood of remission and this raised the possibility that SG2 membership was mediated by sampling diagnoses differences in samples across sites. We thus controlled for schizophrenia diagnosis and repeated the analysis of remission across all timepoints controlling for schizophrenia diagnosis, to reduce the number of comparisons. Results confirmed that remission was associated with higher SG2 membership and lower SG1 membership, negative symptoms, and major depressive disorder diagnoses (Fig. S13). We also controlled for illicit drug use in clinical remission analyses and found similar results (Fig. S13). Finally, we compared the clustering results from the semi-supervised technique (HYDRA) with those form a completely unsupervised technique (k-means + ). Results demonstrated broadly similar brain volumetric results, evidencing considerable stability, except relationships between SG2, education, and remission were not found (Supplementary Materials ‘Investigation of subgroup separation with unsupervised methods’; Fig. S15; Tables S11–S15).

**DISCUSSION**

In this study, we used a large international multi-site heterogeneous sample of individuals with a first-episode of psychosis to validate the presence of the two data-driven neuromorphological subgroups originally derived from a sample of individuals with chronic schizophrenia. We also evaluated whether these subgroups were related to remission over the subsequent course of illness. Our findings show for the first time that these two subgroups [7], SG1 with lower widespread cortical volumes and SG2 with larger striatal volume but otherwise normal brain morphology, are already apparent at the first presentation of illness. Furthermore, our data show a distinct clinical signature separating these subgroups, and that the subgroup presenting with only increased striatal and pallidum volume (SG2) was significantly more likely to achieve remission in subsequent years. These findings support the presence of reproducible neuromorphological subgroups in individuals with psychosis that may help delineate the heterogeneity of brain structure reported in previous neuroimaging research [1, 41, 42]. With further follow-up studies, the findings could also be crucial in informing future research to refine stratified therapeutic approaches and outcome prediction.

The higher proportion of patients with psychosis within the SG1 subgroup relative to HC agrees with our hypothesis based on our previous research in population samples [20]. The findings suggest that the subgroup solutions we identified are not the result of a long duration of illness [7, 8, 10, 11, 13, 14, 16, 17], but are already evident at illness onset, across a wide range of psychosis diagnoses, and before any potential effect of long-term pharmacological treatment. Notably, a similar subgroup to SG1 (but not SG2) has recently also been found in previous unmedicated first-episode schizophrenia samples [16]. In the context of our previous reports of a relationship between SG1, schizophrenia genetic risk, and subthreshold psychosis symptoms [20], the results further point to the presence of a biological vulnerability that may specifically increase risk for the expression of illness in some individuals (i.e., 32% of the FEP sample in this study). Such results support our previous neurodevelopmental hypothesis of SG1 by providing evidence for a potential brain-diathesis model, whereby an existing vulnerability could combine with other risk factors in order to trigger the illness; e.g., family history as found in our multivariable model when compared to SG2 [43] (for further discussion of neurobiological hypotheses see Supplementary Materials). This neurodevelopmental hypothesis was supported by supplementary analyses showing that decreased total brain volume and voxel-wise decreases remain after correction for intracranial volume. The results also contextualize existing neuroimaging research [41] by suggesting that smaller brain volumes are only evident for some individuals with a first episode psychosis.

The equal proportions of FEP patients and HC in SG2 additionally supports our previous population-based research showing the same lack of enrichment [20]. These results suggest that SG2 membership is more likely to be normatively present and does not increase FEP risk. In comparison to SG1 membership, we characterized the sociodemographic and clinical signature to find that SG2 membership was associated with higher education in addition to higher rates of positive symptoms, unemployment, female sex, and chlorpromazine equivalent dose. These results are important as they suggest that, for some individuals in specific brain-based subgroups, the presence of positive symptoms is not necessarily associated with smaller brain volume when examined in the context of the multivariable signature (for example, when higher education and female sex are also considered). Such findings contextualize previous reports of the association between positive symptoms and smaller brain volumes [3] by suggesting that this does not hold for some individuals. It is notable that such a finding has been hypothesized in historical theories of a ‘Type-I’ (non-deficit) schizophrenia suggesting the presence of positive symptoms with less structural brain changes compared to a ‘Type-II’ presentation.
associated with impairment and negative symptoms [44]. In our previous work defining the subgroups [7] we hypothesized that the SG2 subgroup was related to hyperdopaminergic mechanisms, which is now further contextualized by our current results showing a relationship between SG2, female sex, and remission, due to evidence of: increased striatal volumes in remitting females [28, 29], sex differences in striatal dopamine synaptic concentrations and receptors [45], and sex differences in response to antipsychotics [46] (see Supplementary Materials for details).

In univariate tests, remission rates were higher in the SG2 subgroup when timepoints were combined and at 1-year at uncorrected levels. Multivariable remission signatures at 1-year, 5-year, and combined timepoints included SG2 membership combined with decreased schizophrenia diagnosis (and increased diagnoses of schizophrreniform, psychosis not otherwise specified, and brief psychotic disorder). Other notable variables included in the models with SG2 membership were: reduced unemployment (1-year), female sex (5-years), and decreased SG1 membership (combined timepoints). These results support our hypothesis of increased chances of remission in SG2 derived from top-down analysis, was associated with remission, DUP duration of untreated psychosis, PANSS positive and negative syndrome scale, GAF global assessment of functioning, CPZ chlorpromazine equivalent dose, MDD major depressive disorder, Psychosis NOS Psychosis not otherwise specified.

Against the background of the SG1 and SG2 subgroups we identified, it is important to highlight that approximately 34% of FEP individuals did not show either brain signature and were considered to have no subgroup membership (i.e., ‘None’). Multivariable signatures separating this subgroup from the others were also non-significant, potentially supporting the presence of clinical and neuromorphological heterogeneity in presentations. Given that these individuals were also experiencing a psychosis, further research should further investigate this subgroup to potentially detect additional subgroup solutions and finer-grained differences (e.g., using functional MRI or diffusion tensor imaging) that may reveal more precise relationships with symptoms and outcomes in these individuals. In addition, further research in larger samples could consider the SG1 + SG2 subgroup due to its mixed brain signature and higher proportion of individuals with a diagnosis of schizophrenia. However, our study thus highlights how such MRI heterogeneity may influence MRI findings [1, 41, 42] and ultimately obscure biomarker identification in early stage psychosis neuroimaging research [4–6]. Our study thus emphasizes the need for more research on biologically-based individual differences that also accounts for the potential presence.
of premorbid and normatively present brain diversity [7, 8, 10, 11, 13, 14, 17].

STRENGTHS AND LIMITATIONS

This study has multiple strengths as it provides the first replication of the presence of neuromorphological subgroups from a chronic schizophrenia cohort in a large multisite sample of patients at illness onset—a group less confounded by the effects of chronic illness (and long-term treatment)—and in healthy controls. As such, we replicated findings across diverse study protocols and demonstrated the crucial associations between these brain subtypes and clinical variables, suggesting a high degree of generalizability of these subgroups. Furthermore, we also examined the relationship between subgroups across different psychosis diagnoses and with longitudinal outcome data on up to 5-years after the baseline assessment.

However, there are also some important limitations to consider. First, while we believe this has been a strength in generalizability, we should consider the potential effect of differences in MRI acquisition and subject recruitment protocols across sites. We implemented site control protocols that mitigated MRI site variance and controlled for site in clinical analyses, although remaining effects are possible. A related point is that missing variables within each site were addressed using imputation within cross-validation routines during clinical analyses. Nevertheless, these site effects were partly controlled in supplementary analyses that reinforced our finding of a relationship between remission and membership of the SG2 subgroup. Second, although we did not find antipsychotic effects on brain volume or clinical outcomes at baseline, longitudinal treatment could influence remission signatures and needs to be investigated in future research. Third, the baseline and follow-up predictive accuracy related to the clinical associations (e.g., with SG2) was modest (60–69%). Fourth, recreational drug use could be longitudinally investigated to determine interactions with brain signatures and remission. Fifth, the interaction with cognition needs to be considered given the potential relationships with subgroups [20] and remission [50]. Further studies are needed, especially in consortium large samples consisting of clinical high-risk for psychosis groups with homogeneous clinical and cognitive tests, to further clarify these aspects (e.g., PRONIA [39]). Studies in controlled treatment trials could also further investigate differential treatment effects over time.

CONCLUSION

This study validated the presence of specific brain subgroups, originally found in chronic schizophrenia study samples, in a clinically heterogeneous first-episode psychosis sample and demonstrated a significant relationship of these baseline clinical signatures and subsequent symptom remission. The results suggest the possibility of normatively present, premorbid brain types that influence first presentation and outcomes in FEP. Furthermore, they provide an initial, but important, indication that originally found in chronic schizophrenia study samples, in a This study validated the presence of specific brain subgroups, originally found in chronic schizophrenia study samples, in a clinically heterogeneous first-episode psychosis sample and demonstrated a significant relationship of these baseline clinical signatures and subsequent symptom remission. The results suggest the possibility of normatively present, premorbid brain types that influence first presentation and outcomes in FEP. Furthermore, they provide an initial, but important, indication that originally found in chronic schizophrenia study samples, in a clinical context.

CODE AVAILABILITY

HYDRA (https://github.com/Varol/HYDRA) and MIDAS (https://github.com/evorol/MIDAS) code used in this study are publicly available. The machine learning analyses were conducted with NeuroMiner software available at https://github.com/neurominer-glt. Additional code used to assign subgroups and compute signatures is available at https://github.com/ganchand/AJP_Codes.


ACKNOWLEDGEMENTS
Supported by NIH grant R01MH112070 and the National Health and Medical Research Council (NHMRC) grant GA126980. Russel Shinohara was supported by the NIH grant R01MH123550. RCG was supported by the NIH grant R01MH119219.

AUTHOR CONTRIBUTIONS
Conception and design: DBD, GBC, RSK, HGS, EM, GE, SJW, CZ, HS, YF, TMM, REG, RCG, TDS, MVZ, DHW, CP, BCF, GFB, CD, NK, PD. Acquisition of data: DBD, GBC, AP, GE, JD, RSK, HGS, EM, SJW, CZ, HS, YF, MS, PR, PAL, RU, ANK, TMN, BB, REG, MDS, TDS, RMM, MDF, SC, MVZ, DHW, CP, BCF, GFB, CD, NK, PD. Analysis and interpretation of data: DBD, GBC, AP, AK, JW, MA, GH, GE, JD, DS, EV, RSK, HGS, EM, SJW, CZ, AS, RTS, HS, YF, MS, PR, PAL, RU, ANK, TMN, BB, REG, MDS, TDS, RMM, MDF, SC, MVZ, DHW, CP, BCF, GFB, CD, NK, PD. Paper drafting: DBD, GBC, AP, AK, JW, MA, GH, GE, JD, DS, EV, RSK, HGS, EM, SJW, CZ, AS, RTS, HS, YF, MS, PAL, RU, ANK, TMN, BB, REG, MDS, TDR, TDS, MDF, SC, MVZ, DHW, CP, BCF, GFB, CD, NK, PD.

FUNDING
Open Access funding enabled and organized by CAUL and its Member Institutions.

COMPETING INTERESTS
Dr AS has equity in TheraPanacea. Dr PD has received speaking fees from Janssen and Lundbeck. Dr RSK has served as a consultant for Alkermes, Sunovion, and Merck. Dr CP has received funding from the National Health and Medical Research Council, the Lundbeck Foundation, and the Medical Research Future Fund and has received honoraria for talks at educational meetings and as an advisory board member for Lundbeck Australia. Dr RTS has served as a consultant for Octave Bioscience. The other authors report no financial relationships with commercial interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41380-023-02069-0.

Correspondence and requests for materials should be addressed to Dominic B. Dwyer, Nikolaos Koutsouleris or Paola Dazzan.

Reprints and permission information is available at http://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) 2023