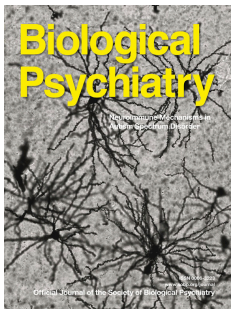
Journal Pre-proof



Neurobiologically Based Stratification of Recent Onset Depression and Psychosis: Identification of Two Distinct Transdiagnostic Phenotypes

Paris Alexandros Lalousis, MSc, Lianne Schmaal, PhD, Stephen J. Wood, PhD, Renate L.E.P. Reniers, PhD, Nicholas M. Barnes, PhD, Katharine Chisholm, PhD, Sian Lowri Griffiths, PhD, Alexandra Stainton, PhD, Junhao Wen, PhD, Gyujoon Hwang, PhD, Christos Davatzikos, PhD, Julian Wenzel, MSc, Lana Kambeitz-Ilankovic, PhD, Christina Andreou, MD, Carolina Bonivento, PhD, Udo Dannlowski, MD, Adele Ferro, PhD, Theresa Liechtenstein, MD, Anita Riecher-Rössler, MD, Georg Romer, MD, Marlene Rosen, PhD, Alessandro Bertolino, MD, Stefan Borgwardt, MD, Paolo Brambilla, MD, Joseph Kambeitz, MD, Rebekka Lencer, MD, PhD, Christos Pantelis, MB BS, MD, MRCPsych, FRANZCP, Stephan Ruhrmann, MD, Raimo K.R. Salokangas, MD, MSc, PhD, PsD, Frauke Schultze-Lutter, PhD, André Schmidt, PhD, Eva Meisenzahl, MD, Nikolaos Koutsouleris, MD, Dominic Dwyer, PhD, Rachel Upthegrove, MBBS FRCPsych, PhD, for the PRONIA Consortium

PII: S0006-3223(22)01156-8

DOI: https://doi.org/10.1016/j.biopsych.2022.03.021

Reference: BPS 14829

To appear in: *Biological Psychiatry*

Received Date: 30 September 2021

Revised Date: 4 February 2022

Accepted Date: 1 March 2022

Please cite this article as: Lalousis P.A., Schmaal L., Wood S.J., Reniers R.L.E.P, Barnes N.M., Chisholm K., Griffiths S.L., Stainton A., Wen J., Hwang G., Davatzikos C., Wenzel J., Kambeitz-Ilankovic L., Andreou C., Bonivento C., Dannlowski U., Ferro A., Liechtenstein T., Riecher-Rössler A., Romer G., Rosen M., Bertolino A., Borgwardt S., Brambilla P., Kambeitz J., Lencer R., Pantelis C., Ruhrmann S., Salokangas R.K.R., Schultze-Lutter F., Schmidt A., Meisenzahl E., Koutsouleris N., Dwyer D., Upthegrove R. & for the PRONIA Consortium, Neurobiologically Based Stratification of Recent Onset Depression and Psychosis: Identification of Two Distinct Transdiagnostic Phenotypes, *Biological Psychiatry* (2022), doi: https://doi.org/10.1016/j.biopsych.2022.03.021.

Neurobiologically Based Stratification of Recent Onset Depression and Psychosis: Identification of Two Distinct Transdiagnostic Phenotypes

Paris Alexandros Lalousis, MSc[1,2]; Lianne Schmaal, PhD[3,4]; Stephen J. Wood, PhD[1,3,4]; Renate L.E.P Reniers, PhD[1,2,5]; Nicholas M. Barnes, PhD[5]; Katharine Chisholm, PhD[1,6]; Sian Lowri Griffiths, PhD[1,2]; Alexandra Stainton, PhD[3,4]; Junhao Wen, PhD[7]; Gyujoon Hwang, PhD[7]; Christos Davatzikos, PhD[7]; Julian Wenzel, MSc[8]; Lana Kambeitz-Ilankovic, PhD[8]; Christina Andreou, MD[9]; Carolina Bonivento, PhD[10,11]; Udo Dannlowski, MD[12]; Adele Ferro, PhD[10,11]; Theresa Liechtenstein, MD[8]; Anita Riecher-Rössler, MD[9]; Georg Romer, MD[10]; Marlene Rosen, PhD[8]; Alessandro Bertolino, MD[13]; Stefan Borgwardt, MD[8,14]; Paolo Brambilla, MD[9,10]; Joseph Kambeitz, MD[15]; Rebekka Lencer, MD, PhD[9,12]; Christos Pantelis, MB BS, MD, MRCPsych, FRANZCP[16]; Stephan Ruhrmann, MD[9]; Raimo K. R. Salokangas, MD, MSc, PhD, PsD[17]; Frauke Schultze-Lutter, PhD[18,19,20]; André Schmidt, PhD[9]; Eva Meisenzahl, MD[18]; Nikolaos Koutsouleris, MD[15]; Dominic Dwyer, PhD[15]*; Rachel Upthegrove, MBBS FRCPsych, PhD[1,2,21*], for the PRONIA Consortium#

* Joint Senior Authors

#See PRONIA consortium author list

Author Affiliations:

1 -Institute for Mental Health, University of Birmingham, Birmingham, United Kingdom

2 -Centre for Human Brain Health, University of Birmingham, Birmingham, United Kingdom

3 -Orygen, Parkville, Australia

4 -Centre for Youth Mental Health, The University of Melbourne, Parkville, Australia

5 -Institute of Clinical Sciences, University of Birmingham, United Kingdom

6 -Department of Psychology, Aston University, United Kingdom

7 -Perelman School of Medicine University of Pennsylvania, United States of America

8 -Department of Psychiatry and Psychotherapy, Faculty of Medicine and University Hospital, University of Cologne, Cologne, Germany

9 -Department of Psychiatry (UPK), University of Basel, Basel, Switzerland

10 -Department of Pathophysiology and Transplantation, University of Milan, Milan, Italy

11 - Department of Neurosciences and Mental Health, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, University of Milan, Milan, Italy

12 –Institute for Translational Psychiatry, University of Münster, Münster, Germany

13 -Department of Basic Medical Sciences, Neuroscience and Sense Organs, University of Bari Aldo Moro, Bari, Italy

14 -Department of Psychiatry and Psychotherapy, University of Lübeck, Lübeck, Germany

15 -Department of Psychiatry and Psychotherapy, Ludwig Maxmilians University, Munich, Germany

16 -Melbourne Neuropsychiatry Centre, University of Melbourne, Melbourne, Australia

17 -Department of Psychiatry, University of Turku, Turku, Finland

18 -Department of Psychiatry and Psychotherapy, University of Düsseldorf, Düsseldorf, Germany

19 -Department of Psychology and Mental Health, Faculty of Psychology, Airlangga University, Surabaya, Indonesia

20 - University Hospital of Child and Adolescent Psychiatry and Psychotherapy, University of Bern, Bern, Switzerland

21 -Birmingham Early Interventions Service, Birmingham Women's and Children's NHS Foundation Trust

Corresponding Author: Paris Alexandros Lalousis, BSc, MSc, School of Psychology, University of Birmingham, 52 Pritchatts Road, Birmingham, B15 2SA, United Kingdom (pal532@student.bham.ac.uk)

Word Count Manuscript

Abstract: 246

Introduction: 719

Methods: 897

Results: 906

Discussion: 1746

Total: 4268 words (excluding abstract)

Tables: 3

Figures: 2

Supplementary Methods and Results: 1 file

Keywords: transdiagnostic, psychosis, depression, clustering, nosology, machine learning

Abstract

Background: Identifying neurobiologically based transdiagnostic categories of depression and psychosis may elucidate heterogeneity, and provide better candidates for predictive modelling. We aimed to identify clusters across patients with recent onset depression (ROD) and recent onset psychosis (ROP) based on structural neuroimaging data. We hypothesized that these transdiagnostic clusters would identify patients with poor outcome and allow more accurate prediction of symptomatic remission than traditional diagnostic structures.

Methods: HYDRA (HeterogeneitY through DiscRiminant Analysis) was trained on whole brain volumetric measures from 577 participants from the discovery sample of the multi-site PRONIA study to identify neurobiologically driven clusters which were then externally validated in the PRONIA replication sample (n=404) and three datasets of chronic samples (COBRE, n=146; MCIC, n=202; MUC, n=470).

Results: The optimal clustering solution was two transdiagnostic clusters (Cluster 1, n=153, 67 ROP, 86 ROD and Cluster 2, n=149, 88 ROP, 61 ROD; ARI=.618). The two clusters contained both ROP and ROD. One cluster had widespread GMV deficits, more positive, negative, and functional deficits (impaired cluster) and one cluster revealed a more preserved neuroanatomical signature and more 'core' depressive symptomatology (preserved cluster). The clustering solution was internally and externally validated and assessed for clinical utility in predicting 9-month symptomatic remission -outperforming traditional diagnostic structures.

Conclusions: We identified two transdiagnostic neuroanatomically informed clusters which are clinically and biologically distinct, challenging current diagnostic boundaries in recent onset mental health disorders. These results may aid understanding of aetiology of poor outcome patients transdiagnostically and improve development of stratified treatments.

Introduction

The current classification of mental disorders is based on a phenomenological approach that uses signs and symptoms to assign a diagnosis. Whilst some diagnoses have high reliability, their usefulness and aetiopathogenetic basis is questionable (1–3). For example, there is considerable commonality of symptoms and neurobiological domains across mental disorders and co-morbidity frequently occurs; with a prevalence of depression in over 40% of people with schizophrenia (4,5) and psychotic symptoms occuring in around 20% of people with depression (6,7).

In terms of brain structure, grey matter volume (GMV) reduction is found in both depression and psychosis, across similar areas such as the anterior insula and the dorsal anterior cingulate cortex (8). This GMV loss has been shown to predate medication exposure, poor functional outcome, neurocognitive deficits, and in the case of clinical high risk for psychosis, transition to frank illness (5,9–11). Symptoms common to depression and schizophrenia, such as social withdrawal, blunted affect, and alogia, are associated with GMV reduction in the cerebellum, while anhedonia and avolition are negatively correlated with left anterior limb of internal capsule white-matter volume (WMV) and positively correlated with left superior longitudinal fasciculus WMV (12).

GMV loss in psychosis and depression may be related to immune dysfunction. Elevated pro-inflammatory cytokines, potentially resulting from genomic predisposition or response to environmental factors, may lead to activation of astrocytic dysfunction and/or microglia activation, resulting in dendritic pruning and synaptic changes (13–15). Indeed, immune dysfunction is implicated in the aetiology of both schizophrenia and depression with cytokines such as IL-6 and CRP detected at elevated levels (16–20), and causality suggested in mendelian randomisation studies of both disorders (17,21).

Currently, diagnoses are not based on underlying brain structure or distinct biological aetiology. Patients whose symptoms are potentially caused by different biological processes may be given the same diagnosis and patients whose symptoms are potentially caused by same biological processes may be provided with a different diagnosis, a practice which may have detrimental effects on outcome prediction development (22–24). Recent research has highlighted this mismatch between diagnostic labels and the clinical and neuroanatomical picture in depression and psychosis (25) and heterogeneity may be particularly pronounced in early stages of developing mental health disorders (26–30). The lack of biological validity of diagnostic groups is thought to be one of the major reasons for poor biomedical translation in psychiatry (31–33).

Only 20% of people with psychosis and 25% of people with depression achieve full remission and response to pharmacological treatment, with the remainder achieving partial response or response without remission (34–37). Biologically-driven illness models, able to relate to those at highest risk of poor outcome and chronicity may allow new and targeted treatments to be delivered early (22). However, recognizing patients on a path to chronic disability, at an early stage, is still difficult in both psychosis and depression (38,39). Previous transdiagnostic research has stressed the need for the use of machine learning (40) and has identified specific patterns of neurocircuit disruption across major psychiatric disorders in emotional reactivity and regulation (41). Reininghaus and colleagues, building on previous calls for a dimensional approach to psychosis (42), have shown the use of multidimensional item response modelling to predict psychosis biotypes transcending traditional diagnostic boundaries; with suggestion of an underlying transdiagnostic dimension across psychotic diagnoses (43–45). Recent semi-supervised machine learning studies using neuroanatomical data have identified the presence of an impaired neuroanatomical cluster which is characterized by overall poorer outcomes and functioning in schizophrenia (46) and

in youth with internalizing symptoms (47). However, there has not yet been a transdiagnostic investigation of neuroanatomy specifically in depression and psychosis.

Herein, we aimed to identify replicable neuroanatomical clusters across patients with recent onset depression (ROD) and recent onset psychosis (ROP). We hypothesized that neuroanatomically derived clusters would be transdiagnostic, and related to distinct phenotypes drawn from symptom, neurocognitive, and inflammatory data across both disorders. We further aimed to explore the predictive validity of neuroanatomically identified clusters and externally validated our neuroanatomically based clusters in chronic depression and chronic schizophrenia, in an accelerated longitudinal design. We also developed supervised machine learning models to predict symptom remission in ROP and ROD and our neuroanatomically based transdiagnostic clusters. We hypothesised that models developed in neuroanatomically based transdiagnostic clusters will show greater predictive accuracy compared to those in traditional diagnostic groups.

Methods

Study design

This study utilizes data from the PRONIA study, an EU-FP7 funded seven-centre study as well as three external validation datasets. Details of the PRONIA study sites, recruitment protocol and quality control procedures can be found in the supplementary methods (1.1, 1.2, 1.3, tables S1, S2, S3) and a prior publication (48). Data used in this analysis included structural MRI, demographic, clinical, neurocognitive and blood-based biomarker measures. See supplement for full details.

Inclusion and Exclusion Criteria

In brief, ROP participants had to meet the following criteria: 1) DSM-IV-TR (Diagnostic and Statistical Manual of Mental Disorders, Text Revision) (49) affective or non-affective psychotic episode (lifetime), 2) criteria for DSM-IV-TR affective or nonaffective psychotic episode fulfilled within past 3 months and 3) onset of psychosis within past 24 months. ROD patients had to meet the following criteria: 1) DSM-IV-TR major depressive episode (lifetime), 2) major depressive disorder criteria fulfilled within past three months and 3) duration of first depressive episode no longer than 24 months. General inclusion criteria can be found in the supplement (1.5).

MRI imaging data acquisition, quality control, and preprocessing

Participants underwent a multi-modal MRI protocol. A minimal harmonization protocol, which the MR sequences across the different scanners had to comply with as well as the imaging preprocessing is described in the supplementary methods (1.3 and 1.4).

Semi-Supervised Machine Learning Analysis:

HeterogeneitY through DiscRiminant Analysis (HYDRA) (50) is a semi-supervised machine learning clustering algorithm able to dissect disease heterogeneity by portioning patients based on patterns or transformations between the sub-populations (i.e., clusters) from the patient group and the reference group (i.e., healthy controls) through the use of a convex polytope formed by combination of multiple linear max-margin classifiers (i.e., support vector machines) and is able to regress out nuisance covariates, such as age and gender. We used the python version of HYDRA (https://github.com/anbai106/pyHYDRA) (50) to simultaneously classify patients (ROP+ROD) from HC, and partition patients into clusters based on disease-related heterogeneity using structural MRI.

ComBat Harmonization

To mitigate site effects, prior to applying HYDRA, the R version of the ComBat harmonization technique was employed (https://github.com/Jfortin1/ComBatHarmonization). ComBat utilizes an empirical Bayesian framework that removes variance which is attributed to scanner differences while retaining disease effects. To further ensure that disease variance would be retained distinct from scanner variance, ComBat was trained on HC and then derived estimates were applied to the patients.

Model Training

We used whole volume (GMV and cerebrospinal fluid) brain measures derived from 280 regions of the neuromorphometrics atlas parcellation (CAT12) (four regions excluded due to zero variance) from 577 participants with ROP and ROD and HC (discovery sample of the PRONIA study). ROP patients and ROD patients were grouped together into one patient group. HYDRA was trained using a repeated hold-out cross-validation strategy (i.e., 1000 repetitions with 80% of the data for training in each repetition). Age, sex, and Total Intracranial Volume (TIV) were controlled as covariates. HYDRA was ran for 2 to 8 clustering solutions, and Adjusted Rand Index (ARI) was used to measure cluster stability. The most stable cluster solution was selected for further analysis. The statistical significance of clusters was assessed in three ways including testing our clustering solution against a gaussian distribution which assumes a dimensional severity explanation of our data. Details can be found in the supplement (1.11).

Phenotype Characterization

Identified clusters were compared to each other and to HC in terms of neurocognitive performance, blood-based biomarker (IL1ra, S100B, IL6, TNFα, CRP, TGFβ, and BDNF) (see supplement 1.6) and symptom differences (PANSS, BDI, SANS) with univariate statistics corrected for multiple comparisons using false-discovery rate (FDR).

Neuroanatomical differences were examined using voxel-based morphometry (two sample t-test, SPM12), to identify the brain regions that the neuroanatomically derived clusters differed on. See supplementary material (section 1.14) for further granular investigation of clinical and inflammatory marker differences between clusters.

Independent and External Validation

To examine the generalizability of neuroanatomically based clusters we developed a SVM model, using the 280 features that our HYDRA model was trained on (46), to classify patients from the discovery sample into the identified clusters. This SVM was applied to the PRONIA-independent replication sample of ROP and ROD patients (N=404), collected at a different timescale from the discovery sample (May 2016 to February 2019). ComBat was trained on the replication HC and applied to the replication transdiagnostic patient group to mitigate site effects in the replication dataset. The SVM validation model that was trained on the discovery data was then applied to the replication data.

We externally validated the neuroanatomically based PRONIA clusters using the developed SVM model, in three MRI datasets of patients with chronic schizophrenia (Centre for Biomedical Research Excellence (COBRE) and Mind Clinical Imaging Consortium (MCIC) and chronic depression (Munich (MUC)) in an accelerated longitudinal design framework (see supplementary methods 1.9 and 1.10).

Predictive Utility

We trained SVM models using symptom and blood-based biomarker data to predict symptom recovery (as defined by a Global Assessment of Functioning-Symptom (GAF-S) score of ≥61) (51) at 9 months. To assess the predictive utility within the neuroanatomically based clusters and within ROP and ROD groups we trained 4 different SVM models (one for each different diagnosis ROP/ROD/Cluster 1/Cluster 2) and compared their predictive accuracy in

terms of area under the ROC curve, balanced accuracy, sensitivity and specificity. Details can be found in the supplement (1.8). A detailed figure of the analysis pipeline can be seen in figure 1.

Results:

Demographic Information

One hundred fifty-five participants with ROP, 147 patients with ROD, and 275 HC from the discovery sample were included in the HYDRA semi-supervised machine learning analysis. The mean age of the ROP group was 25.3 [SD 5.5], the mean age of the ROD group was 25.9 [SD 6.2]), and the mean age of the HC was 25.5 [SD 6.4]. The ROP group consisted of 96 male and 59 female patients, the ROD group had 66 male and 81 female patients, and the HC group had 107 male and 168 female participants. A summary of sociodemographic and clinical information is provided in table 1. Sociodemographic and clinical information for the PRONIA replication and external validation samples (COBRE, MCIC, and MUC) is provided in the supplement (1.9).

HYDRA Semi-Supervised Machine Learning Analysis

The optimal clustering solution was two transdiagnostic clusters (Cluster 1, n=153, 67 ROP, 86 ROD and Cluster 2, n=149, 88 ROP, 61 ROD, ARI: .618). Patients in cluster 1 had a mean age of 26.2 [6.2] and the ones in Cluster 2 had a mean age of 24.9 [5.4]. There were 78 male and 75 female patients in cluster 1 and 84 male and 65 female patients in cluster 2. The two clusters did not differ in terms of age (p=.071), sex distribution (p=.358), total intracranial volume (p=.144), or medication exposure and differed in terms of original diagnosis distribution (p=.008). A sociodemographic and clinical description of the two clusters can be found in table 1.

Cluster Statistical Significance

The clusters were statistically significant 1) in terms of whether they would be different than if there was no disease related variability present (p=.010), 2) in terms of whether the disease structures were different (p<.001), and 3) in terms of whether the data could be better explained by a single Gaussian distribution (p<.001) suggesting that our data could not be explained in terms of a single Gaussian (continuous) distribution assuming a dimensional severity model. Details of the statistical significance tests can be found in the supplement (1.11).

Clinical Characteristics Associated with Neuroanatomically based clusters

Cluster 2 revealed a more severe symptom presentation compared to cluster 1 with significantly higher scores in the positive (t(287)=-2.8, p=.020), negative (t(287)=-2.2, p=.040), and general (t(287)=-2.7, p=.010) PANSS domains. Patients in cluster 2 had higher negative symptoms in SANS symptoms of affective flattening (t(284)=-2.7, p=.010), alogia(t(282)=-3.0, p=.020), and attention deficit (t(255)=-2.2, p=.040). Patients in Cluster 2 also showed worse functioning (Global Functioning-Role) (t(291)=-2.3, p=.030). There were no statistically significant differences between the two clusters in terms of neurocognition or blood-based biomarker data in univariate analysis. All p values have been fdr corrected. See supplement tables S5, S6, and S7. In supplementary multivariate SVM analysis our neuroanatomically based clusters were separable using cognitive data (BAC: 56.6%, sensitivity: 57.5%, specificity: 55.7%, AUC: 0.58, p=0.01). Patients in cluster 2 mainly exhibited worse cognitive performance in a visual recognition and recall task (Rey–Osterrieth complex figure) and patients in cluster 1 mainly performed worse in verbal memory tasks (Rey Auditory Verbal Learning Test) (See supplementary figures S6, S7, and S8). The two clusters were also separable (BAC: 58.7%, sensitivity: 54.9%, specificity: 62.4%, AUC: 0.59, p=0.01) in blood-based biomarkers, with patients in cluster 2 having elevated levels of CRP and TNFα (See supplementary figures S9, S10, and S11).

VBM analysis of neuroanatomically based clusters

We conducted a VBM analysis for the purpose of demonstrating the brain regions that the two clusters differed in. Here, Cluster 2 exhibited widespread GMV loss compared to Cluster 1 and also compared to HC in areas including the Superior Temporal Gyrus, the Cingulate Gyrus, and the Thalamus among others. Cluster 1 revealed increased GMV compared to HC in cerebellar areas. These results can be seen in figure 2 and in the supplement (tables S7 and S8 and figure S2).

Independent and External Validation

In independent validation the two-cluster model showed generalisability in the PRONIA replication sample with patients classified into the two clusters in the replication sample showing similar clinical and neuroanatomical patterns to the ones from the discovery sample (supplement section 1.18). When externally applied to the MCIC and COBRE (chronic schizophrenia) and MUC (chronic depression), patients from datasets with a higher mean of age and/or longer duration of illness were, more often placed in Cluster 2 as indicated by negative decision scores. The effects of duration of illness and age were statistically significant, $F(2,278) = 27.88$, $p<.001$. Post hoc analyses using the Tukey HSD post hoc criterion for significance indicated that the mean decision score was significantly lower in the MUC group compared to the MCIC ($p<.001$). Mean decision score differences between the MCIC and COBRE ($p=.078$) showed a trend towards statistical significance. The results can be seen in table 2.

Prognostic Validation

Within the neuroanatomically based clusters, stacking a blood-based biomarker (IL1ra, CRP, TNFα, BDNF, and TGFβ) SVM model to a symptom data (baseline PANSS, BDI, and GAF-S individual item scores) SVM model (i.e., a combined model) increased accuracy for

predicting symptomatic recovery at 9 months (GAF-S) with BAC of 71.2% for cluster 1 and 57.0% for cluster 2. This outperformed a similar stacked blood-based biomarker and symptom data SVM model predicting GAF-S in ROP and ROD groups (table 3). A Kruskal-Wallis H test showed that there is a statistically significant difference between the outer cross-validation folds (CV2) BAC of the different models $H(3)=22.9$, $p<0.001$. Post-hoc Mann-Whitney U test results can be found in the supplement (1.13).

Discussion

In this study, we identified two transdiagnostic clusters across psychosis and depression, using semi-supervised machine learning and neuroanatomical data in a large sample of recent onset depression and psychosis patients. Both clusters contained similar numbers of patients with depression and psychosis, however they were clinically distinct, with one cluster being characterized by more general and negative symptom loading and functional impairment, widespread GMV loss, (hereafter called the “impaired” cluster) and one cluster characterized by fewer symptoms, less GMV loss, and less functional impairment but more ‘core’ depressive symptomatology (hereafter called the “preserved” cluster). The neuroanatomically based clusters were generalizable to a replication sample and further externally validated in three datasets of patients with chronic illness. Patients with chronic illness, with a higher duration of illness and mean age, were more likely to be classified into the impaired cluster. We were further able to demonstrate that SVM learning models using clinical and blood-based biomarker data to predict symptom remission at 9 months showed a higher accuracy in the neuroanatomically derived clusters compared to traditional diagnostic categories.

The precise aetiology of mental illnesses including psychosis and depression remains elusive despite decades of research, with a stagnation in advance of new pharmacological and psychotherapeutic treatments (52–54). Our results suggest that current diagnostic categories,

particularly in early stages of illness, may mask transdiagnostic phenotypes which include an identifiable group with greater impairment and poorer chance of remission across disorders. In our impaired cluster, patients had reduced GMV in areas that have been identified as central to the disease processes of both schizophrenia and depression, such as the superior temporal gyrus, the anterior cingulate, the insula, and the thalamus (55–58). In our analysis, a significant number of patients with depression, who may be perceived as having a less severe illness and better prognostic outlook than patients with psychosis, were ascribed to the impaired phenotype, suggesting that they are on a path towards poor outcome. Conversely, a significant number of patients with psychosis were not assigned to the impaired group, and therefore potentially have an identifiable early signature of good prognosis, which was further indicated by the fact that predicting 9-month symptomatic outcomes in that group was more accurate than traditional diagnostic groupings.

Categorical diagnoses have survived because some individuals (specifically those with chronic established illness) do indeed fit within these nosological entities and more valid solutions remain elusive to date (59). However, within the scope of affective and non-affective major psychiatric diseases, the Kraepelinian dichotomy of dementia praecox and manic-depressive psychosis has long been challenged. Studies have shown that our understanding of the clinical and neurobiological distinction between disorders may be particularly challenging during early phases of illness (5,25,60,61). The concept of affective disorders as a differential diagnosis for psychosis, particularly in the early years of illness is waning, with recent research suggesting a central and causal role for depression in the pathogenesis of psychosis and mutual biological underpinnings. This further challenges the distinction between affective and non-affective pathways to psychosis (25,61–63). Fischer and Carpenter (64) suggest that reducing heterogeneity in syndromes is essential to decisively address the Kraepelinian dichotomy. Despite the fact that dementia praecox does not directly

map to non-affective psychosis, the Verrücktheit (chronic non-affective psychoses) made distinct in Kraepelin's first Edition (1883) led to the (mis)understanding that schizophrenia was non-affective (65). The impaired cluster which contains both patients with schizophrenia and depression has more cognitive symptoms and a brain signature that is identified in our chronic replication sample. Deficit schizophrenia is a concept introduced over 30 years ago to reduce clinical heterogeneity and suggests the existence of a homogeneous schizophrenia subtype with persistent trait negative symptoms (66). The impaired cluster we identified could be characterized as a transdiagnostic deficit cluster across depression and psychosis due to its higher load of negative symptoms, a previously proposed marker of the deficit syndrome across diagnoses (67). Furthermore, our findings of greater GMV reduction in the impaired cluster corroborate previous research which identified temporal GMV reduction as a marker of very poor outcome (68). Our neuroanatomically derived clusters contained both patients with depression and psychosis in recent onset, replicated in our independent PRONIA sample. This suggests lack of diagnostic hierarchy across depression and psychosis, and that some syndromes may hold equal weight in association with poor outcome regardless of relationship to diagnosis. These results add to the challenge of the separation between affective and non-affective psychoses with affective and psychotic diagnostic groups featuring in both clusters; corroborating previous studies which found that high affective symptom scores were equally common in patients with affective and non-affective psychosis and question the clinical validity of such a distinction (69).

Our results support the common biological susceptibility model of psychiatric disorders and suggest that the biological underpinnings of disease course, at least in depression and psychosis, may be related to transdiagnostic mechanisms, which are potentially hidden by current nosological systems. A similar transdiagnostic model has previously been reported in genomic research, which has shown a certain degree of overlap in the biological

susceptibility to mental illness across mood and psychotic disorders; evidence of a transdiagnostic biological cause of major psychiatric disorders is evident with the identification of genetic variants that confer a transdiagnostic risk for bipolar, major depressive disorder, and schizophrenia related to the Major Histocompatibility Complex featuring in both schizophrenia and depression genome wide association studies (70,71). Our finding that elevated pro-inflammatory cytokines add to predictive accuracy of poor outcome in an impaired phenotype suggest that this genomic immune influence may be ongoing in those on a path to poor outcomes. Schizophrenia GMV deficits in the hippocampus, temporal gyrus, and cerebellum are associated with genetic factors such as SATB2, GABBR2, and CACNA1C (72). A common genetic basis between risk for altered brain structure and neuro-psychiatric disorders has been conferred by findings of risk variant enrichment associations with brain structural phenotypes across diagnoses (73). Our results suggest a transdiagnostic cluster of GMV impairment suggestive of common biological underpinnings for poor outcome across depression and psychosis with potentially more valid structures than traditional diagnostic categories for use in predicting symptomatic remission.

Heterogeneity and co-morbidity may be especially pronounced in the early stages of these disorders; this creates diagnostic uncertainty and difficulties in predicting disease and treatment course (26–30). Our results suggest that a bottom-up approach based on neurobiological data may be more reliable in the elucidation of patients with potential for greater impairment and offer a potential future solution for the diagnostic challenges of mental illness. Our external validation findings show that the impaired cluster potentially identifies patients who are on a path to chronic illness from early stages of illness, given that the majority of patients in the external validation sample with chronic illness fell into the same cluster as our impaired group. This has potentially significant clinical implications in terms of personalised treatment and focused recovery interventions. The fact that patients

from chronic samples with a higher mean age and illness duration were more likely to be assigned to the impaired cluster could be an indication that our neuroanatomically based clusters identify an accelerated transdiagnostic brain aging effect in recent onset samples, corroborating previous brain age studies (74,75).

Strengths and Limitations

The present analysis includes several strengths including a large dataset with rich clinical, neurocognitive, biomarker, and imaging data from both recent onset psychosis and depression groups, independent and external validation, as well as significance testing of our clustering solutions (e.g. by testing whether the data could be better explained by a Gaussian distribution which assumes a dimensional severity explanation of the data). Furthermore, the technique we used for the identification of subgroups (HYDRA), offers a solution to issues that are usually associated with clustering based on unsupervised machine learning models which are built on biological data such as the detection of groups that may reflect underlying nuisance variance such as age, gender, body type, and common ancestry (genetics) (76). Nevertheless, our results should be interpreted with caution as there are certain limitations. Due to the nature of our recent onset sample and using a healthy control sample as a reference group in the semi-supervised model, there is a risk that the differences between the groups are not as marked as would be seen in more chronic cases. We addressed that limitation by performing permutation tests to robustly assess the significance of the identified clusters. Furthermore, our models were developed in recent onset patients with a significantly lower mean age than that of our external validation samples. We addressed that limitation by following a robust pipeline that removed the age and site effects while retaining the disease variance in the data. Although we developed an accelerated longitudinal design with the use of recent onset and chronic samples and had a 9 month follow-up for prediction of symptom remission, definitive findings would need large longitudinal datasets with repeated measures,

such as functional outcome, over many years. Finally, we only used neuroanatomical features to parse neurobiological variance among complex clinical presentations. Psychiatric illness is not a single variable problem and we have addressed that by examining whether the brain-based clustering solution is reflected in the phenotypic, cognitive, and inflammatory levels. Future studies should consider using multiple biological measures and larger population-level data to encompass the pleiomorphic nature of clinical entities such as depression and psychosis.

Conclusions

Using semi-supervised machine learning, we were able to identify two neuroanatomically based transdiagnostic clusters. One cluster was characterized by an impaired functional and neurocognitive profile and greater symptomatic loading and GMV loss while the other cluster was characterized by a more preserved neuroanatomical and reduced symptom signature. Our distinct impaired cluster included patients with depression and psychosis and may provide insight into transdiagnostic aetiopathogenetic pathways of chronicity and poor outcome. The identified clusters have been derived in recent onset samples using structural MRI and could eventually lead to the development of MRI-based prediction and decision-making tools. In external validation, older patients with longer duration of schizophrenia and depression were assigned in the impaired cluster suggesting a potential identifiable transdiagnostic signature of chronicity and path to poor outcome at the early disease stages. Using clinical and blood-based biomarker data, we were able to predict symptomatic and functional remission more accurately in the derived clusters compared to traditional diagnostic groups. Whilst such challenge to current diagnostic structures will need significant further replication and longer follow-up, identifying a transdiagnostic signature of poor prognosis has the potential to aid new and targeted treatment strategies across early stages of mental disorder.

Acknowledgements:

Author Contributions: Mr. Lalousis, Prof. Koutsouleris, Prof. Upthegrove, and Dr Dwyer had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. All authors reviewed, revised, and approved the final version of the manuscript.

Acquisition and analysis of data: Lalousis, Wood, Chisholm, Griffiths, Stainton Borgwardt, Brambilla, Pantelis, Andreou, Dannlowski, Riechler-Rössler, Romer, Bonivento, Dwyer, Ferro, Liechtenstein, Rosen, Schmidt, Koutsouleris, and Upthegrove

Drafting of the manuscript: Lalousis, Schmaal, Reniers, Koutsouleris, Upthegrove, and Dwyer

Critical revision of the manuscript for important intellectual content: Lalousis, Schmaal, Wood, Barnes, Griffiths, Stainton, Wen, Hwang, Davatzikos, Bertolino, Borgwardt, Brambilla, Kambeitz, Lencer, Pantelis, Ruhrmann, Salokangas, Schultze-Lutter, Dwyer, Schmidt, Meisenzahl, Koutsouleris, and Upthegrove

Statistical analysis and interpretation of data: Lalousis, Schmaal, Upthegrove, and Dwyer

Obtained funding and designed the study: Wood, Bertolino, Borgwardt, Brambilla, Kambeitz, Lencer, Pantelis, Ruhrmann, Salokangas, Meisenzahl, Koutsouleris, and Upthegrove

Administrative, technical, or material support: Chisholm, Borgwardt, Brambilla, Liechtenstein, Rosen, Schmidt, Meisenzahl, Koutsouleris, and Upthegrove

Supervision: Wood, Schmaal, Reniers, Borgwardt, Brambilla, Schultze-Lutter, Koutsouleris, Upthegrove, and Dwyer

Funding: The PRONIA study is a Collaboration Project funded by the European Union under the 7th Framework Programme under grant agreement n° 602152.

#The PRONIA consortium:

The authors listed here performed the screening, recruitment, rating, examination, and follow-up of the study participants. They were involved in implementing the examination protocols of the study, setting up its IT infrastructure, and organizing the flow and quality

control of the data analyzed in this manuscript between the local study sites and the central study database.

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

Linda Betz, Anne Erkens, Eva Gussmann, Shalaila Haas, Alkomiet Hasan, Claudius Hoff, Ifrah Khanyaree, Aylin Melo, Susanna Muckenhuber-Sternbauer, Janis Köhler, Ömer Öztürk, Nora Penzel, David Popovic, Adrian Rangnick, Sebastian von Saldern, Rachele Sanfelici, Moritz Spangemacher, Ana Tupac, Maria Fernanda Urquijo, Johanna Weiske, Antonia Wosgien

Department of Psychiatry and Psychotherapy, University of Cologne, Cologne, Germany

Karsten Blume, Dominika Gebhardt, Nathalie Kaiser, Ruth Milz, Alexandra Nikolaides, Mauro Seves, Silke Vent, Martina Wassen

Department of Psychiatry (Psychiatric University Hospital, UPK), University of Basel, Switzerland

Christina Andreou, Laura Egloff, Fabienne Harrisberger, Claudia Lenz, Letizia Leanza, Amatya Mackintosh, Renata Smieskova, Erich Studerus, Anna Walter, Sonja Widmayer

Institute for Mental Health & School of Psychology, University of Birmingham, United Kingdom

Chris Day, Mariam Iqbal, Mirabel Pelton, Pavan Mallikarjun, Alexandra Stainton, Ashleigh Lin

Department of Psychiatry, University of Turku, Finland

Alexander Denissoff, Anu Ellilä, Tiina From, Markus Heinimaa, Tuula Ilonen, Päivi Jalo, Heikki Laurikainen, Antti Luutonen, Akseli Mäkela, Janina Paju, Henri Pesonen, Reetta-Liina Säilä, Anna Toivonen, Otto Turtonen

General Electric Global Research Inc., USA

Ana Beatriz Solana, Manuela Abraham, Nicolas Hehn, Timo Schirmer

Workgroup of Paolo Brambilla, University of Milan, Italy:

Department of Neuroscience and Mental Health, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, University of Milan, Milan, Italy: Carlo Altamura, Marika Belleri, Francesca Bottinelli, Adele Ferro, Marta Re

Programma2000, Niguarda Hospital, Milan: Emiliano Monzani, Maurizio Sberna

San Paolo Hospital, Milan: Armando D'Agostino, Lorenzo Del Fabro

Villa San Benedetto Menni, Albese con Cassano (CO): Giampaolo Perna, Maria Nobile, Alessandra Alciati

Workgroup of Paolo Brambilla at the University of Udine, Italy

Department of Medical Area, University of Udine, Udine, Italy: Matteo Balestrieri, Carolina Bonivento, Giuseppe Cabras, Franco Fabbro

IRCCS Scientific Institute "E. Medea", Polo FVG, Udine: Marco Garzitto, Sara Piccin

Disclosures

Pantelis has participated on Advisory Boards for Janssen-Cilag, Astra-Zeneca, Lundbeck, and Servier. He has received honoraria for talks presented at educational meetings organised by Astra-Zeneca, Janssen-Cilag, Eli-Lilly, Pfizer, Lundbeck and Shire. Koutsouleris received honoraria for talks presented at education meetings organized by Otsuka/Lundbeck. Upthegrove reports grants from the Medical Research Council, grants from National Institute for Health Research: Health Technology Assessment, grants from European Commission - Research: The Seventh Framework Programme, and personal speaker fees from Sunovion, outside the submitted work. All other authors report no biomedical financial interests or potential conflicts of interest.

	ROP Group	ROD Group	t/χ2	P Value	Cluster 1 (Preserved)	Cluster 2 (Impaired)	t/z/χ2	P Value	HC	HC vs Impaired Cluster	t/z/χ2	P Value	HC vs Preserved Cluster	t/z/χ2	P Value
Sample Sizes, No.	155	147			153	149			275						
Original Diagnostic Group (ROP/ROD)					(67/86) 43.2%/58.5%	(88/61) 56.8%/41.5%	χ2 = 7.04	.008							
Age, Mean (SD)	25.3 (5.5)	25.9 (6.2)	*t* = -.879	.380	26.2 (6.2)	24.9 (5.4)	*t* = 1.81	.071	25.5 (6.4)	*t* = .887	.375		*t* = -1.035	.301	
Sex (Male/Female)	96/59	66/81	χ^2 = 8.8	.003	78/75	84/65	χ^2 = .88	.358	107/168	χ^2 = 11.9	.001		χ^2 = 5.8	.016	
Total Intracranial Volume, Mean (SD)	1531.6 (141.9)	1500.6 (144.3)	*t* = 1.87	.061	1504.6 (144.0)	1528.7 (142.8)	*t* = -1.46	.144	1518.5 (140.8)	*t* = -.708	.481		*t* = .975	.330	
Medication, Mean Cumulative Sum (SD)															
CPZE					5122.7 (16501.2)	11191.7 (52988.6)	*t* = -1.24	.214							
OLAE					390.5 (1780.0)	173.9 (551.4)	*t* = -1.32	.187							
SSRIE					3095.7 (10409.5)	2504.3 (7975.8)	*t* = .510	.610							
BENZOE					282.8 (1031.5)	578.6 (3625.2)	*t* = -.888	.375							

SCID Diagnosis, No. (%)								
Schizophrenia	63 (40.6)	0 (0)			22 (14.4)	41 (27.5)		
Schizophreniform Disorder	12 (7.7)	0 (0)			3 (2.0)	9 (6.0)		
Schizoaffective Disorder	8 (5.2)	0 (0)			4 (2.6)	4 (2.7)		
Delusional Disorder	8 (5.2)	0 (0)			7 (4.6)	1 (0.7)		
Psychotic Disorder NOS	22 (14.2)	0 (0)			11 (7.2)	11 (7.4)		
Major Depressive Disorder	13 (8.4)	140 (95.2)			88 (57.5)	65 (43.6)		
Bipolar Disorder I	9 (5.8)	0 (0)			4 (2.6)	5 (3.4)		
Other	20 (12.9)	7 (4.8)			14 (9.1)	13 (8.7)		
PANSS Positive Mean (SD)	17.5 (6.3)	7.6 (1.2)	t = 18.25	<.001	11.5 (5.8)	13.1 (7.4)	t = -2.83	.02
PANSS Negative Mean (SD)	16.4 (7.9)	12.2 (4.7)	t = 5.43	<.001	13.5 (6.3)	15.2 (7.2)	t = -2.21	.04
PANSS General Mean (SD)	35.7 (11.6)	27.1 (6.5)	t = 7.99	<.001	29.8 (8.2)	33.0 (11.4)	t = -2.71	.01

Table 1.Sample Sociodemographics. Sample Sizes, Participants per Study Site, Age, Sex, Total Intracranial Volume, Medication. (Abbreviations: ROP=Recent Onset Psychosis, ROD=Recent Onset Depression, HC=Healthy Controls, SD=Standard Deviation, CPZE=Chlorpromazine Equivalent, OLAE=Olanzapine Equivalent, SSRIE=Selective Serotonin Reuptake Inhibitor Equivalent, BENZOE=Benzodiazepine Equivalent, SCID=Structured Clinical Interview for DSM Disorders, NOS=Not Otherwise Specified, PANSS=Positive and Negative Symptom Scale)

	COBRE	MCIC	MUC
Diagnosis	Schizophrenia	Schizophrenia	Depression
Sample Size	71	107	103
Age M (SD)	38.1 (13.9)	34.5 (11.1)	42.1 (11.9)
Duration of Illness M (SD)	16.8 (12.9)	10.9 (10.9)	5.8 (7.7)
Mean Decision Score	-.04 (.63)	.15 (.71)	-.47 (.48)

Table 2. External validation results. Decisions scores reflect mean distance of patients from the hyperplane separating the two clusters. Positive decision scores indicate assignment to cluster 1 (preserved cluster) and negative decision scores indicate assignment to cluster 2 (impaired cluster). $F(2,278) = 27.88$, $p<.001$

	True Positive, No.	True Negative, No.	False Positive, No.	False Negative, No.	Correct Classification Rate Unremitted, %	Correct Classification Rate, Remitted, %	Balanced Accuracy, %	Positive Predictive Value, %	Negative Predictive Value, %	AUC	Model P Value
Stacked ROP 9-month Model	20	33	19	29	40.8	63.5	52.1	51.3	53.2	0.56	0.38
Stacked ROD 9-month Model	53	11	13	26	67.1	45.8	56.5	80.3	29.7	0.54	0.17
Stacked Preserved Cluster 9-month Model	19	54	11	13	59.4	83.1	71.2	63.3	80.6	0.72	0.07
Stacked Impaired 9-month Model	35	25	16	31	53.0	61.0	57.0	68.6	44.6	0.58	0.18

Table 3. SVM models predicting 9-month GAF-S remission. H(3)=22.9, p<0.001.

References

1. Cuthbert BN (2014): The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry* 13: 28–35.
2. First MB, Rebello TJ, Keeley JW, Bhargava R, Dai Y, Kulygina M, *et al.* (2018): Do mental health professionals use diagnostic classifications the way we think they do? A global survey. *World Psychiatry* 17: 187–195.
3. Fusar-Poli P, Solmi M, Brondino N, Davies C, Chae C, Politi P, *et al.* (2019): Transdiagnostic psychiatry: a systematic review. *World Psychiatry* 18: 192–207.
4. Conley RR, Ascher-Svanum H, Zhu B, Faries D, Kinon BJ (2007): The Burden of Depressive Symptoms in the Long-Term Treatment of Patients With Schizophrenia. *Schizophr Res* 90: 186–197.
5. Upthegrove R, Marwaha S, Birchwood M (2017): Depression and Schizophrenia: Cause, Consequence, or Trans-diagnostic Issue? *Schizophr Bull* 43: 240–244.
6. Johnson J, Horwath E, Weissman MM (1991): The Validity of Major Depression With Psychotic Features Based on a Community Study. *Arch Gen Psychiatry* 48: 1075–1081.
7. Ohayon MM, Schatzberg AF (2002): Prevalence of Depressive Episodes With Psychotic Features in the General Population. *AJP* 159: 1855–1861.
8. Goodkind M, Eickhoff SB, Oathes DJ, Jiang Y, Chang A, Jones-Hagata LB, *et al.* (2015): Identification of a Common Neurobiological Substrate for Mental Illness. *JAMA Psychiatry* 72: 305–315.
9. Meisenzahl EM, Seifert D, Bottlender R, Teipel S, Zetzsche T, Jäger M, *et al.* (2010): Differences in hippocampal volume between major depression and schizophrenia: a comparative neuroimaging study. *Eur Arch Psychiatry Clin Neurosci* 260: 127–137.
10. Pantelis C, Yücel M, Bora E, Fornito A, Testa R, Brewer WJ, *et al.* (2009): Neurobiological Markers of Illness Onset in Psychosis and Schizophrenia: The Search for a Moving Target. *Neuropsychology Review* 19: 385–398.

11. Reniers RLEP, Lin A, Yung AR, Koutsouleris N, Nelson B, Cropley VL, *et al.* (2017): Neuroanatomical Predictors of Functional Outcome in Individuals at Ultra-High Risk for Psychosis. *Schizophrenia Bulletin* 43: 449–458.

12. Chuang J-Y, Murray GK, Metastasio A, Segarra N, Tait R, Spencer J, *et al.* (2014): Brain structural signatures of negative symptoms in depression and schizophrenia. *Front Psychiatry* 5: 116.

13. Corsi-Zuelli F, Deakin B (2021): Impaired regulatory T cell control of astroglial overdrive and microglial pruning in schizophrenia. *Neuroscience & Biobehavioral Reviews* 125: 637–653.

14. Laskaris L, Mancuso S, Shannon Weickert C, Zalesky A, Chana G, Wannan C, *et al.* (2021): Brain morphology is differentially impacted by peripheral cytokines in schizophrenia-spectrum disorder. *Brain Behav Immun* 95: 299–309.

15. Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, *et al.* (2016): Schizophrenia risk from complex variation of complement component 4. *Nature* 530: 177–183.

16. Khandaker GM, Pearson RM, Zammit S, Lewis G, Jones PB (2014): Association of Serum Interleukin 6 and C-Reactive Protein in Childhood With Depression and Psychosis in Young Adult Life. *JAMA Psychiatry* 71: 1121–1128.

17. Khandaker GM, Cousins L, Deakin J, Lennox BR, Yolken R, Jones PB (2015): Inflammation and immunity in schizophrenia: implications for pathophysiology and treatment. *Lancet Psychiatry* 2: 258–270.

18. Noto C, Maes M, Ota VK, Teixeira AL, Bressan RA, Gadelha A, Brietzke E (2015): High predictive value of immune-inflammatory biomarkers for schizophrenia diagnosis and association with treatment resistance. *The World Journal of Biological Psychiatry* 16: 422–429.

19. Osimo EF, Pillinger T, Rodriguez IM, Khandaker GM, Pariante CM, Howes OD (2020): Inflammatory markers in depression: A meta-analysis of mean differences and variability in 5,166 patients and 5,083 controls. *Brain, Behavior, and Immunity* 87: 901–909.

20. Upthegrove R, Manzanares-Teson N, Barnes NM (2014): Cytokine function in medication-naive first episode psychosis: A systematic review and meta-analysis. *Schizophrenia Research* 155: 101–108.

21. Perry BI, Upthegrove R, Kappelmann N, Jones PB, Burgess S, Khandaker GM (2021): Associations of immunological proteins/traits with schizophrenia, major depression and bipolar disorder: A bi-directional two-sample mendelian randomization study. *Brain, Behavior, and Immunity*. https://doi.org/10.1016/j.bbi.2021.07.009

22. Linden DEJ (2012): The Challenges and Promise of Neuroimaging in Psychiatry. *Neuron* 73: 8–22.

23. Stephan KE, Binder EB, Breakspear M, Dayan P, Johnstone EC, Meyer-Lindenberg A, *et al.* (2016): Charting the landscape of priority problems in psychiatry, part 2: pathogenesis and aetiology. *Lancet Psychiatry* 3: 84–90.

24. Stephan KE, Bach DR, Fletcher PC, Flint J, Frank MJ, Friston KJ, *et al.* (2016): Charting the landscape of priority problems in psychiatry, part 1: classification and diagnosis. *Lancet Psychiatry* 3: 77–83.

25. Lalousis PA, Wood SJ, Schmaal L, Chisholm K, Griffiths SL, Reniers RLEP, *et al.* (2021): Heterogeneity and Classification of Recent Onset Psychosis and Depression: A Multimodal Machine Learning Approach. *Schizophrenia Bulletin*. https://doi.org/10.1093/schbul/sbaa185

26. Baca-Garcia E, Perez-Rodriguez MM, Basurte-Villamor I, Fernandez del Moral AL, Jimenez-Arriero MA, Gonzalez de Rivera JL, *et al.* (2007): Diagnostic stability of psychiatric disorders in clinical practice. *Br J Psychiatry* 190: 210–216.

27. Keshavan MS, Brady R (2011): Biomarkers in schizophrenia: we need to rebuild the Titanic. *World Psychiatry* 10: 35–36.

28. Koutsouleris N, Meisenzahl EM, Borgwardt S, Riecher-Rössler A, Frodl T, Kambeitz J, *et al.* (2015): Individualized differential diagnosis of schizophrenia and mood disorders using neuroanatomical biomarkers. *Brain* 138: 2059–2073.

29. Pope MA, Joober R, Malla AK (2013): Diagnostic stability of first-episode psychotic disorders and persistence of comorbid psychiatric disorders over 1 year. *Can J Psychiatry* 58: 588–594.

30. Salvatore P, Baldessarini RJ, Khalsa H-MK, Amore M, Vittorio CD, Ferraro G, *et al.* (2013): Predicting Diagnostic Change Among Patients Diagnosed With First-Episode DSM-IV-TR Major Depressive Disorder With Psychotic Features. *J Clin Psychiatry* 74: 723–731.

31. Brückl TM, Spoormaker VI, Sämann PG, Brem A-K, Henco L, Czamara D, *et al.* (2020): The biological classification of mental disorders (BeCOME) study: a protocol for an observational deep-phenotyping study for the identification of biological subtypes. *BMC Psychiatry* 20: 213.

32. Kapur S, Phillips AG, Insel TR (2012): Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry* 17: 1174–1179.

33. Kendler KS (2013): What psychiatric genetics has taught us about the nature of psychiatric illness and what is left to learn. *Mol Psychiatry* 18: 1058–1066.

34. Brown S, Kim M, Mitchell C, Inskip H (2010): Twenty-five year mortality of a community cohort with schizophrenia. *Br J Psychiatry* 196: 116–121.

35. Health (UK) NCC for M (2014): *PSYCHOSIS AND SCHIZOPHRENIA IN ADULTS*. National Institute for Health and Care Excellence (UK). Retrieved February 26, 2020, from https://www.ncbi.nlm.nih.gov/books/NBK333029/

36. Kern DM, Cepeda MS, Defalco F, Etropolski M (2020): Treatment patterns and sequences of pharmacotherapy for patients diagnosed with depression in the United States: 2014 through 2019. *BMC Psychiatry* 20. https://doi.org/10.1186/s12888-019-2418-7

37. Owen MJ, Sawa A, Mortensen PB (2016): Schizophrenia. *The Lancet* 388: 86–97.

38. Dinga R, Marquand AF, Veltman DJ, Beekman ATF, Schoevers RA, van Hemert AM, *et al.* (2018): Predicting the naturalistic course of depression from a wide range of clinical, psychological, and biological data: a machine learning approach [no. 1]. *Translational Psychiatry* 8: 1–11.

39. Klosterkötter J, SCHULTZE-LUTTER F, BECHDOLF A, RUHRMANN S (2011): Prediction and prevention of schizophrenia: what has been achieved and where to go next? *World Psychiatry* 10: 165–174.

40. Fusar-Poli P, Stringer D, M. S. Durieux A, Rutigliano G, Bonoldi I, De Micheli A, Stahl D (2019): Clinical-learning versus machine-learning for transdiagnostic prediction of psychosis onset in individuals at-risk. *Transl Psychiatry* 9: 259.

41. McTeague LM, Rosenberg BM, Lopez JW, Carreon DM, Huemer J, Jiang Y, *et al.* (2020): Identification of Common Neural Circuit Disruptions in Emotional Processing Across Psychiatric Disorders. *Am J Psychiatry* 177: 411–421.

42. Maj M (2016): The need for a conceptual framework in psychiatry acknowledging complexity while avoiding defeatism. *World Psychiatry* 15: 1–2.

43. Tamminga CA, Ivleva EI, Keshavan MS, Pearlson GD, Clementz BA, Witte B, *et al.* (2013): Clinical phenotypes of psychosis in the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP). *Am J Psychiatry* 170: 1263–1274.

44. Reininghaus U, Böhnke JR, Chavez-Baldini U, Gibbons R, Ivleva E, Clementz BA, *et al.* (2019): Transdiagnostic dimensions of psychosis in the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP). *World Psychiatry* 18: 67–76.

45. Quattrone D, Di Forti M, Gayer-Anderson C, Ferraro L, Jongsma HE, Tripoli G, *et al.* (2019): Transdiagnostic dimensions of psychopathology at first episode psychosis: findings from the multinational EU-GEI study. *Psychological Medicine* 49: 1378–1391.

46. Chand GB, Dwyer DB, Erus G, Sotiras A, Varol E, Srinivasan D, *et al.* (2020): Two distinct neuroanatomical subtypes of schizophrenia revealed using machine learning. *Brain* 143: 1027–1038.

47. Kaczkurkin AN, Sotiras A, Baller EB, Barzilay R, Calkins ME, Chand GB, *et al.* (2020): Neurostructural Heterogeneity in Youth with Internalizing Symptoms. *Biol Psychiatry* 87: 473–482.

48. Koutsouleris N, Kambeitz-Ilankovic L, Ruhrmann S, Rosen M, Ruef A, Dwyer DB, *et al.* (2018): Prediction Models of Functional Outcomes for Individuals in the Clinical High-Risk State for Psychosis or With Recent-Onset Depression. *JAMA Psychiatry* 75: 1156–1172.

49. American Psychiatric Association (Ed.) (1994): *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV ; Includes ICD-9-CM Codes Effective 1. Oct. 96*, 4. ed., 7. print. Washington, DC.

50. Varol E, Sotiras A, Davatzikos C (2017): HYDRA: revealing Heterogeneity of imaging and genetic patterns through a multiple max-margin Discriminative Analysis framework. *Neuroimage* 145: 346–364.

51. *Recovery from Schizophrenia: An International Perspective: A Report from the WHO Collaborative Project, the International Study of Schizophrenia* (2007): New York, NY, US: Oxford University Press, pp xix, 370.

52. Brand SJ, Möller M, Harvey BH (2015): A Review of Biomarkers in Mood and Psychotic Disorders: A Dissection of Clinical vs. Preclinical Correlates. *Curr Neuropharmacol* 13: 324–368.

53. Dean J, Keshavan M (2017): The neurobiology of depression: An integrated view. *Asian J Psychiatr* 27: 101–111.

54. Radua J, Ramella-Cravaro V, Ioannidis JPA, Reichenberg A, Phiphopthatsanee N, Amir T, *et al.* (2018): What causes psychosis? An umbrella review of risk and protective factors. *World Psychiatry* 17: 49–66.

55. Gray JP, Müller VI, Eickhoff SB, Fox PT (2020): Multimodal Abnormalities of Brain Structure and Function in Major Depressive Disorder: A Meta-Analysis of Neuroimaging Studies. *AJP* 177: 422–434.

56. Kempton MJ, Salvador Z, Munafò MR, Geddes JR, Simmons A, Frangou S, Williams SCR (2011): Structural Neuroimaging Studies in Major Depressive Disorder: Meta-analysis and Comparison With Bipolar Disorder. *Arch Gen Psychiatry* 68: 675.

57. Schmaal L, Hibar DP, Sämann PG, Hall GB, Baune BT, Jahanshad N, *et al.* (2017): Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. *Mol Psychiatry* 22: 900–909.

58. van Erp TGM, Walton E, Hibar DP, Schmaal L, Jiang W, Glahn DC, *et al.* (2018): Cortical brain abnormalities in 4474 individuals with schizophrenia and 5098 controls via the ENIGMA consortium. *Biol Psychiatry* 84: 644–654.

59. Potuzak M, Ravichandran C, Lewandowski KE, Ongür D, Cohen BM (2012): Categorical vs dimensional classifications of psychotic disorders. *Compr Psychiatry* 53: 1118–1129.

60. Birchwood M, Iqbal Z, Upthegrove R (2005): Psychological pathways to depression in schizophrenia: studies in acute psychosis, post psychotic depression and auditory hallucinations. *Eur Arch Psychiatry Clin Neurosci* 255: 202–212.

61. Craddock N, Owen MJ (2010): The Kraepelinian dichotomy – going, going... but still not gone. *Br J Psychiatry* 196: 92–95.

62. Upthegrove R, Lalousis P, Mallikarjun P, Chisholm K, Griffiths SL, Iqbal M, *et al.* (n.d.): The Psychopathology and Neuroanatomical Markers of Depression in Early Psychosis. *Schizophr Bull*. https://doi.org/10.1093/schbul/sbaa094

63. Herniman SE, Phillips LJ, Wood SJ, Cotton SM, Liemburg EJ, Allott KA (2021): Interrelationships between depressive symptoms and positive and negative symptoms of recent onset schizophrenia spectrum disorders: A network analytical approach. *Journal of Psychiatric Research* 140: 373–380.

64. Fischer BA, Carpenter WT (2009): Will The Kraepelinian Dichotomy Survive DSM-V? *Neuropsychopharmacology* 34: 2081–2087.

65. Kendler KS (2020): The Development of Kraepelin's Concept of Dementia Praecox: A Close Reading of Relevant Texts. *JAMA Psychiatry* 77: 1181–1187.

66. Carpenter WT, Heinrichs DW, Wagman AM (1988): Deficit and nondeficit forms of schizophrenia: The concept. *The American Journal of Psychiatry* 145: 578–583.

67. Peralta V, Cuesta MJ (2003): The nosology of psychotic disorders: a comparison among competing classification systems. *Schizophr Bull* 29: 413–425.

68. Mitelman SA, Buchsbaum MS (2007): Very poor outcome schizophrenia: Clinical and neuroimaging aspects. *Int Rev Psychiatry* 19: 345–357.

69. van Os J, Gilvarry C, Bale R, van Horn E, Tattan T, White I, Murray R (2000): Diagnostic value of the DSM and ICD categories of psychosis: an evidence-based approach. *Social Psychiatry and Psychiatric Epidemiology* 35: 305–311.

70. Li H, Chang H, Song X, Liu W, Li L, Wang L, *et al.* (2019): Integrative analyses of major histocompatibility complex loci in the genome-wide association studies of major depressive disorder. *Neuropsychopharmacol* 44: 1552–1561.

71. Mokhtari R, Lachman HM (2016): The Major Histocompatibility Complex (MHC) in Schizophrenia: A Review. *J Clin Cell Immunol* 7: 479.

72. Luo N, Tian L, Calhoun VD, Chen J, Lin D, Vergara VM, *et al.* (2019): Brain function, structure and genomic data are linked but show different sensitivity to duration of illness and disease stage in schizophrenia. *NeuroImage: Clinical* 23: 101887.

73. Mufford MS, Stein DJ, Dalvie S, Groenewold NA, Thompson PM, Jahanshad N (2017): Neuroimaging genomics in psychiatry—a translational approach. *Genome Medicine* 9: 102.

74. Han LKM, Dinga R, Hahn T, Ching CRK, Eyler LT, Aftanas L, *et al.* (2020): Brain aging in major depressive disorder: results from the ENIGMA major depressive disorder working group. *Mol Psychiatry* 1–16.

75. Schnack HG, van Haren NEM, Nieuwenhuis M, Hulshoff Pol HE, Cahn W, Kahn RS (2016): Accelerated Brain Aging in Schizophrenia: A Longitudinal Pattern Recognition Study. *Am J Psychiatry* 173: 607–616.

76. Dinga R, Schmaal L, Penninx BWJH, van Tol MJ, Veltman DJ, van Velzen L, *et al.* (2019): Evaluating the evidence for biotypes of depression: Methodological replication and extension of. *NeuroImage: Clinical* 22: 101796.

Legends

Legend Figure 1: Analysis Pipeline Overview. This figure provides an overview of the analysis pipeline undertaken in this study. ROP and ROD patients were combined into one transdiagnostic group and ComBat was trained on HC and applied to the patients in order to remove site related variance from the data. The HC and the patient data were then entered into the HYDRA algorithm with age, sex, and TIV, added as covariates. HYDRA was trained using a repeated hold-out cross-validation strategy (i.e., 1000 repetitions with 80% of the data for training in each repetition). The clusters were validated in the PRONIA replication sample and the three external datasets. Identified clusters were assessed for statistical significance and were then analyzed for clinical and VBM differences. Furthermore, the predictive utility of the clusters was assessed.

Legend Figure 2: Impaired Cluster (Cluster 2) GMV Reductions Compared to the Preserved Cluster (Cluster 1). GMV reductions are observed in the Middle Frontal Gyrus, Superior Frontal Gyrus, Superior Temporal Gyrus, Medial Frontal Gyrus, Cingulate Gyrus, Right Cerebellum, Left Cerebellum, Precuneus, Precentral Gyrus, Inferior Frontal Gyrus, Anterior Cingulate, Insula, Parahippocampal Gyrus, Left Fusiform Gyrus, Hippocampus, Lingual Gyrus, Amygdala, Thalamus, Cuneus, Middle Occipital Gyrus, Right Fusiform Gyrus,

Inferior Temporal Gyrus, and Middle Temporal Gyrus. Peak voxel MNI coordinates can be found in the supplement (table s7).

Legend Table 1: Sample Sociodemographics. Sample Sizes, Participants per Study Site, Age, Sex, Total Intracranial Volume, Medication. (Abbreviations: ROP=Recent Onset Psychosis, ROD=Recent Onset Depression, HC=Healthy Controls, SD=Standard Deviation, CPZE=Chlorpromazine Equivalent, OLAE=Olanzapine Equivalent, SSRIE=Selective Serotonin Reuptake Inhibitor Equivalent, BENZOE=Benzodiazepine Equivalent, SCID=Structured Clinical Interview for DSM Disorders, NOS=Not Otherwise Specified, PANSS=Positive and Negative Symptom Scale)

Legend Table 2: External validation results. Negative decision scores indicate assignment to cluster 2 (impaired cluster). $F(2,278) = 27.88$, $p<.001$

Legend Table 3: SVM models predicting 9-month GAF-S remission. $H(3)=22.9$, $p<0.001$.

155 ROP

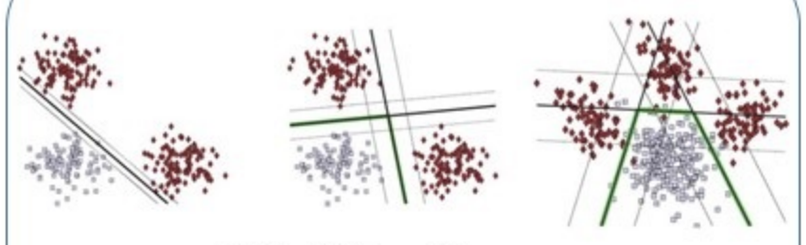
147 ROD

275 HC

ComBat

302 Transdiagnostic Patients

HYDRA



- 1000 CV Repetitions
- 2-8 Clustering Solutions
- Age, Sex, TIV Covariates

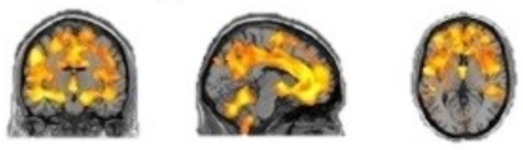
Significance Testing

- Feature Shuffling = $p < 0.001$
- SigClust = $p < 0.001$

- Two Clusters, ARI: .618
- Preserved Cluster (N=153), Impaired Cluster (N=149)
- Preserved Cluster: 67 ROP/86 ROD
- Impaired Cluster: 88 ROP/61 ROD

VBM And Clinical Analysis

Impaired Cluster:



- Changes in appetite and sleeping patterns
- Suspiciousness/persecution and hostility
- Anxiety and disturbance of volition
- Alogia
- Grooming and hygiene
- Social inattentiveness
- Attention
- Global Functioning
- Elevated CRP and TNFα

Prognostic Validation

SVM 9 Month GAF-S Prediction

Clinical Model Features:

- PANSS Item Scores
- BDI Item Scores
- GAF-S Baseline

Inflammation Model Features:

- IL1ra, CRP, TNFα, BDNF, TGFβ, age, sex, tobacco use, and BMI

Stacking BAC:
ROP Model: 62.3%
ROD Model: 56.1%
Impaired Cluster Model: 61.7%
Preserved Cluster Model: 72%

Independent Validation

Replication Sample

152 HC

ComBat

252 TR PAT

Replication Sample

SVM Trained on Original Data

External Validation

HC Subsample from PRONIA, MCIC, MUC, and COBRE (n = 549)

ComBat

Pooled HC and Patients from PRONIA, MCIC, and COBRE (n = 1391)

Dynamic Standardization

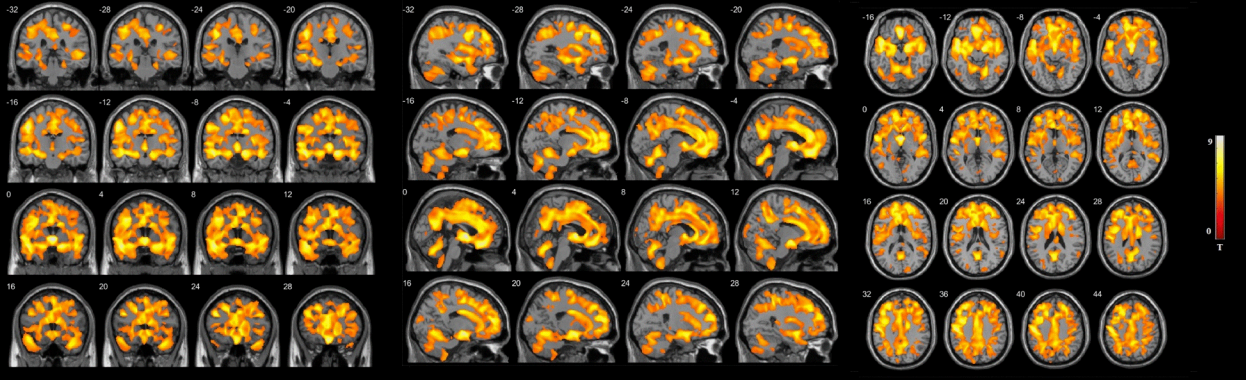
a) For each ComBat corrected patient identified a ComBat corrected HC within age window of +/- 5 years

b) Computed median and std at the feature level

c) Z-Standardized patients' data

SVM Model trained on Standardized Data

Applied to COBRE, MCIC, MUC


-32
-28
-24
-20
-16
-12
-8
-4
0
4
8
12
16
20
24
28
32
36
40
44
9
0
T