Title

Schizophrenia polygenic risk scores, clinical variables and genetic pathways as predictors of phenotypic traits of bipolar I disorder

*Maria Grigoroiu-Serbanescu [1], *Tracey van der Veen [2], Tim Bigdeli [3], Stefan Herms [4, 5], Carmen C Diaconu [6], Ana Iulia Neagu [6], Nicholas Bass [2], Johan Thygesen [2, 7], Andreas J Forstner [5], Markus M. Nöthen [5], Andrew McQuillin [2]

[1] Psychiatric Genetics Research Unit, Alexandru Obregia Clinical Psychiatric Hospital, Bucharest, Romania;

[2]Molecular Psychiatry Laboratory, Division of Psychiatry, University College London, London, UK;

[3]SUNY Downstate Medical Center, Brooklyn, New York, USA;

[4] Department of Biomedicine, University of Basel, Basel, Switzerland;

[5] Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Germany;

[6] Stefan S Nicolau Institute of Virology, Bucharest, Romania

[7] Institute of Health Informatics, University College London, London, UK.

*Equal contribution

Corresponding author: Maria Grigoroiu-Serbanescu (e-mail:maria.serbanescu@gmail.com)

Abstract

Word count = 246

Aim. We investigated the predictive value of polygenic risk scores (PRS) derived from the schizophrenia GWAS (Trubetskoy et al., 2022) **(SCZ3)** for phenotypic traits of bipolar disorder type-I (BP-I) in 1878 BP-I cases and 2751 controls from Romania and UK.

Methods: We used PRSice-v2.3.3 and PRS-CS for computing SCZ3-PRS for testing the predictive power of SCZ3-PRS alone and in combination with clinical variables for several BP-I subphenotypes and for pathway analysis. Non-linear predictive models were also used.

Results. SCZ3-PRS significantly predicted psychosis, incongruent and congruent psychosis, general age-of-onset (AO) of BP-I, AO-depression, AO-Mania, rapid cycling in univariate regressions. A negative correlation between the number of depressive episodes and psychosis, mainly

incongruent and an inverse relationship between increased SCZ3-SNP loading and BP-I-rapid cycling were observed. In random forest models comparing the predictive power of SCZ3-PRS alone and in combination with nine clinical variables, the best predictions were provided by combinations of SCZ3-PRS-CS and clinical variables closely followed by models containing only clinical variables. SCZ3-PRS performed worst. Twenty-two significant pathways underlying psychosis were identified.

Limitations. The combined RO-UK sample had a certain degree of heterogeneity of the BP-I severity: only the RO sample and partially the UK sample included hospitalized BP-I cases. The hospitalization is an indicator of illness severity. Not all UK subjects had complete subphenotype information.

Conclusion. Our study shows that the SCZ3-PRS have a modest clinical value for predicting phenotypic traits of BP-I. For clinical use their best performance is in combination with clinical variables.

Keywords: schizophrenia polygenic score, bipolar disorder subphenotypes, psychosis, individual pathway analysis

Highlights:

- Schizophrenia polygenic risk score derived from the schizophrenia GWAS (Trubetskoy et al., 2022) (SCZ3-PRS) significantly predicted psychosis, incongruent and congruent psychosis, general age-of-onset (AO) of BP-I, AO-depression, AO-mania, and rapid cycling in univariate regressions.
- The best associations of SCZ3-PRS were with psychosis, incongruent psychosis, and age of onset of depression in two PRS computation methods (PRS-CS and threshold and clumping method).
- A negative correlation between the number of depressive episodes and psychosis and an inverse relationship between SCZ3-SNP loading and BP-I rapid cycling were observed.
- The best predictions of BP-I subphenotypes were provided by combinations of SCZ3-PRS-CS and clinical variables, while SCZ3-PRS alone generated the worst predictions in machine learning models.
- Pathway analysis of psychosis in BP-I identified 22 genetic pathways with the highest associations for ***ZNF318***, ***Apoptosis***, and ***Mitochondrion***.

YELLOW = CORRECTED TEXT

Word count main text = 5157 (5151 without chapter titles)

1. Introduction

Developmental psychopathology, family, and genetic studies showed shared liability to schizophrenia (SCZ) and bipolar disorder (BP), as well as a certain degree of specificity that represents the basis for distinct diagnoses. Previous GWAS evidenced a genetic correlation of 0.70 between SCZ and BP based on common single nucleotide polymorphisms (SNPs) (Cross-Disorder Group of PGC, 2019). An analysis of the most recent GWAS of 11 major psychiatric disorders (Grotzinger et al., 2022) identified four genomic factors summarizing their overlapping genetic architecture. One factor was the "psychotic factor**"** clustering common SNPs for BP and SCZ.

Although many SNPs contribute to both disorders, SNP subsets have a larger effect in SCZ than in BP and vice-versa (Ruderfer et al., 2018), which may contribute to clinical specificity. Associations were also described between SCZ-SNP-sets derived from the SCZ-GWAS-2014 (Ripke et al., 2014) and some phenotypic traits of BP such as age of onset (AO) (Ruderfer et al., 2018; Kalman et al., 2021), psychosis (Leonenko et al., 2018; Calafato et al., 2018; Stahl et al., 2019), although these associations were not always replicated (Kalman et al., 2019**).** Stahl et al. (2019) showed that the SCZ polygenic risk score (PRS) loading is higher in BP-type I (BP-I) than in BP-II and Allardyce et al., (2018**)** found that the SCZ-PRS (2014 set) loading was higher in BP-I with psychosis than in BP-I without psychosis. Treatment response in BP was also shown to be influenced by the SCZ-PRS (2014 set) (International Consortium on Lithium Genetics, 2018; Schubert et al., 2021).

A problem with large-scale GWAS samples is their phenotypic heterogeneity generated by the population heterogeneity, the different diagnostic criteria and recruitment settings, and inclusion of several BP-types, usually BP-I, BP-II, schizoaffective patients (e.g. Kalman et al., 2019; Stahl et al., 2019).

The objective of our study was to investigate the predictive value of PRS derived from the PGC Schizophrenia GWAS 2022 (Trubetskoy et al., 2022) (SCZ3) for phenotypic traits of BP-I in two **phenotypically homogeneous and well characterized samples**: a Romanian (RO) sample and an UK sample.

2. Methods

The sample used in our investigation comprised 1878 BP-I cases and 2751 controls from Romania and UK. All participants in this study were of European ancestry and provided written

informed consent. The study was approved by the ethical committees of the two participating centers.

2.1. *Clinical Diagnosis of the Romanian BP-I sample and controls*

All patients and controls were of Romanian descent. **Genealogical information** about parents and all four grandparents was obtained through direct interview of the subjects in order to ensure a sample with a genetic population structure as homogeneous as possible. Unrelated BP-I patients (N=574) were recruited from consecutive admissions in the Obregia Psychiatric Hospital of Bucharest. Patients already with an ICD-10 diagnosis of BP assigned by the treating psychiatrist were investigated in **face-to-face interviews during the last days of hospitalization** with the Diagnostic Interview for Genetic Studies (DIGS) (NIMH version 1999) (elaborated by Nurnberger et al., 1994) based on DSM-IV-criteria (APA, 1994). The interview was conducted **by a researcher with long clinical experience** in order to confirm the BP-I subtype and exclude **residual mood incongruent psychotic symptoms during remissions** for avoiding inclusion of schizoaffective patients. This information was confirmed by first degree relatives for 64% of the cases. The diagnosis of BP-I was assigned according to DSM-IV-criteria based on combined information from DIGS, medical records (3 to 38 hospitalizations depending on patient) and patient relatives.

The illness AO was defined as the age at which the proband first met DSM-IV criteria for a manic, mixed, or major depressive episode.

Population-based controls from Romania **(**N=534) were screened with the DIGS for a lifetime history of any major affective or schizoaffective disorders, schizophrenia and any other psychosis, obsessive compulsive disorder, eating disorders, alcohol or drug addiction. Unaffected controls were included in the study. (**details in the Supplement).**

2.2. *Clinical diagnosis of the UK BP-I sample and controls*

1304 BP-I subject**s** were given a NHS clinical diagnosis of ICD-10 bipolar disorder and fulfilled Research Diagnostic Criteria for BP-I with clinical data collected by Schizophrenia and Affective Disorder Schedule-lifetime (SADS-L) (Endicott & Spitzer, 1978). The participants were also rated with the 90-item OPCRIT (McGuffin et al., 1991). In the UK sample AO was defined as the age at which the proband first met criteria for a manic, mixed, or major depressive episode according to item 4 from OPCRIT.

The majority of the population-based UK controls (N=2,217) were screened with the SADS-L for lifetime history of any major affective or schizoaffective disorders, schizophrenia and any other psychosis, obsessive compulsive disorder, eating disorders, alcohol or drug addiction. Only

unaffected controls were included in the study. A small part of the UK controls were random blood donors not screened for psychiatric disorders **(N=384; 14.8%).**

The two samples shared clinical (**Table_S1**) and genetic data (**Tables_S2-S4**) facilitating their comparison. There were similar proportions of males and females across the national samples **(Table_S1)** and across cases and controls (BP-I cases : 718 males, 1160 females; controls: 993 males, 1758 females) (χ2=1.296, P=0.254).

Since the mean and median AO differed between samples (Z=2.661, P=0.007) the AO was normalised by rank-based inverse-normal transformation .

2.3. *Genotyping and imputation in the Romanian sample*.

The BP-I cases and controls were genome-wide genotyped on Illumina Omni Express or GSAMD beadchips at the Institute of Human Genetics, Bonn, Germany. Stringent quality control was applied to the genotype information. Individuals were excluded on the basis of having incorrect sex assignments; excessive heterozygosity or homozygosity (more than 10 sd above the mean) (1 patient excluded); missing 5% or more of their genotypes and evidence of relatedness (pi_hat>0.2). SNPs were excluded if they had a minor allele frequency (MAF) <0.1% , deviated substantially from the Hardy-Weinberg equilibrium ($P<10^{-6}$) or had high missingness rate (> 0.05).

QC-ed genotypes were phased using EAGLE (v2.3.5) (Loh et al., 2016) and imputed to the 1000 Genomes Project (Phase 3.v5) using Minimac3 (v2.0.1) (Das et al., 2016) with default parameters. The genotyped variants were aligned to the human reference genome (GRCh37). Non-complementary SNPs with info/rsq>0.8, MAF>1% in the Romanian sample were chosen followed by LD clumping (500 SNP window, 100 SNP overlap, rsq<0.05) based on 1000Genomes to remove SNPs with rsq>0.1 within 1 Mb range.

Principal component analysis (PCA) was performed with EIGENSTRAT v6.1.4 (Price et al., 2006) for all of the 1000Genomes Project subjects and for only the European 1000 Genomes subjects excluding Finnish samples. The lambda based on the Romanian sample was 1.011 and lambda1000 was 1.034, thus indicating a relatively homogenous population with very little evidence of substructure (**PCA description in Supplement).**

2.4. *Genotyping and imputation, in the UK sample*

UK sample was genotyped with either the Illumina PsychArray beadchip or Illumina Global Screening Array (GSA) at the Broad Institute, MA, US. Standard QC was performed as described fully in a previous publication (Grigoroiu-Serbanescu et al., 2020). Individuals were removed if their genotyped sex differed from the reported sex. **SNPs were included if : SNPs had an INFO score >**

0.8, a MAF>0.01, SNP missingness < 0.05; subject missingness<0.02 (after sample removal); autosomal heterozygosity deviation (| Fhet | <0.2); difference in SNP missingness between cases and controls < 0.02; and SNP Hardy-Weinberg equilibrium (P >10−6) in cases and controls.

Imputation was performed using the prephasing/imputation stepwise approach implemented in Eagle v2.3.5 ((Loh et al, 2016)) and Minimac3 (Das et al., 2016**)** to the Haplotype Reference Consortium (HRC) reference panel v1.0 (Loh et al., 2016). Post imputation QC excluded all SNPs with INFO score (R^2) <0.9. We retained the most significant SNPs within a physical distance < 250 kb and an R^2>0.1, based on 1000Genomes Project EUR reference data (since our samples are of European ancestry). Relatedness checks excluded subjects with pi_hat>0.2. PCA was performed after imputation with EIGENSTRATv6.1.4 (Price et al., 2006). The genomic inflation factor for the UK cohort was lambda=1.004, and lambda1000=1.009, suggesting low risk from confounding factors, e.g. population stratification or cryptic relatedness, in our polygenic signal.

2.5. *Polygenic risk score computation in the Romanian and UK samples*

We constructed individual-level PRS from imputed dosages by two methods: clumping and threshold (C+T) and continuous shrinking (PRS-CS). (**explanation in Supplement**).

C+T (clumping and association threshold) method

The GWAS summary statistics for schizophrenia (Trubetskoy et al., 2022) (SCZ3) were obtained from the Psychiatric Genomics Consortium (PGC). A Romanian sample was not included in this GWAS and the UK (UCL) sample was excluded. Individual-level PRS for SCZ3 were calculated in each sample using the C+T method implemented in PRSice v2.3.3 (Euesden et al., 2015; Choi and O'Reilly, 2019) using its default settings.

Normalized PRS for eight P-thresholds (PT) between $P=5x10^{-8}$ and P=0.05 were computed using the overlapping SNPs between the discovery SCZ3 European set (7,627,438) and SNPs in our samples weighting the sum of SNPs associated at a certain PT by their effect size estimates. In the RO sample 3,917,045 SNPs (99,194 SNPs with P-values≤0.05) overlapped with the SCZ3-SNP set, while in the UK sample 4,698,650 SNPs overlapped with the SCZ3 summary statistics. R (R-Core-Team, 2020) was used to standardize the PRS scores (mean=0, SD=1), to facilitate correct interpretation of the individual relative risk (Pain et al., 2022).

PRS-CS

PRSs were also calculated with weights estimated via continuous-shrinkage PRS (PRS-CS) (Ge et al., 2019) software, using a Bayesian regression framework that adjusts SNP effect sizes via a

continuous shrinkage by multivariate modelling of the genetic architecture. The PRS-CS auto-setting was selected, allowing the algorithm to learn the global shrinkage (tuning) parameter from the discovery GWAS data, a pseudo-validation method which performs well comparable to a grid search with tuning parameters (Pain et al., 2021**)**. Default settings were used for all other PRS-CS parameters. The GWAS summary statistics for SCZ3 (Trubetskoy et al., 2022) and European LD reference panel generated from the 1000 Genomes Project were used. The PLINK v2.0 (Chang et al., 2015) *score* function was used to produce raw PRS scores from the posterior means of the estimated SNP effects from PRS-CS to calculate individual-level SCZ3-PRS. The PRS for each sample were standardized to mean=0, sd=1 with R.

2.6. Pathway-specific PRS analysis

Pathway-specific PRS were applied to all patients with the PRS estimated separately in each cohort using PRSet (Choi et al., 2023). This provided an individual-level representation of genetic burden for SCZ3 in BP-I within a gene-set, in contrast to other gene-set methods, e.g., MAGMA (de Leeuw et al., 2015), which can only estimate the association between a gene set and the phenotype at the population level.

In the current study, the permutation test calculated the individual-level gene set PRS. PRSet, is a group of Gene Set Analysis (GSA) methods based on PRS. GSA is sometimes interchangeably called "pathway analyses". PRSet calculated gene-set PRS to study the aetiology at the pathway level**.**

Pathway PRSs were computed for 4629 subjects to evaluate the association between genomic pathways and risk for psychosis in BP-I. There were 1,321 BP-I cases with psychosis (congruent and incongruent), 533 BP-I cases without psychosis and 2,751 controls. 14 UK cases had no psychosis data. Each pathway PRS was tested for association by regressing the phenotype against the pathway PRS. Additionally, PRSet analysis evaluated pathway enrichment by computing a "competitive" enrichment P-value using a permutation method. This accounts for the number of SNPs in pathways, independent of the same SNP across pathways. Gene-sets (N=31,927) were obtained from the MSigDB database using GSEA software (http://software.broadinstitute.org/gsea/msigdb/index.jsp). Competitive P-values were calculated using 10,000 permutations and a threshold of ≤ 1 was applied to ensure enough SNPs within pathways (Choi et al., 2023). Permutation testing considers the correlation structure in the genome which other multiple correction tests do not (Joo et al., 2016) **(**details in the supplement**).**

2.7. *Statistical methods*

SCZ3-PRS performance to predict BP-I subphenotypes, was assessed in linear and logistic regression models. Variance explained in dichotomous outcomes were expressed as Nagelkerke R^2, converted to the liability scale to account for sample prevalences (Lee et al., 2012), using a BP-I population prevalence of 1% (Merikangas et al., 2011; Humpston et al., 2021). Random forest (RF) models were employed to extend BP-I subphenotype predictions to include clinical as well as genetic data. RMSE (Root Mean Squared Error) or AUC (Area Under the ROC curve) were used to gauge prediction accuracy, as well as model fit. **Elastic net penalty regression** (cv.glmnet in R) effect sizes (logOR), were provided to check the robustness of the RF models' variable importance rankings (Mean Decreased Accuracy, MDA), to assess the predictive power of ten subphenotypes, when classifying psychosis and its subtypes (incongruence and congruence). The lowest level of the discriminative value of AUC is 0.71 and an AUC ≥0.79 is considered strongly discriminative (de Hond et al., 2022). For having clinical utility AUC and the positive prediction value (PPV) must reach at least 0.8 (McMahon, 2014). **(details in Supplement**).

The false discovery rate (FDR) (Benjamini and Hochberg, 1995) was used to correct for multiple testing. All statistical analyses were conducted using R.4.2.2. (R-Core-Team, 2020).

3. Results

We analysed the predictive power of SCZ3-PRS-CS (main text tables) and of eight P-thresholds (PT) containing SNPs significantly associated with SCZ ($P=10^{-8}$ to $P=0.05$) (supplementary tables) for several phenotypic traits of BP-I (general AO, AO-first depression, AO-first mania, psychosis, congruent and incongruent psychosis) in the RO, the UK and the combined RO-UK samples.

3.1. *Differentiation of cases and controls*.

First, we tested the ability of the SCZ3-PRS-CS and of the eight SCZ3-PRS-PTs to differentiate the cases from controls in each national sample and in the combined RO-UK sample. All versions of SCZ3-PRS distinguished the cases from controls at highly significant corrected P-values (**Table 1, Tables_S5-S6**) .

3.2. *Phenotypic traits of bipolar I disorder*

Subsequently we analysed several phenotypic traits of BP-I: the general age of onset (AO) of BP-I irrespective of polarity at onset, AO of the first depressive episode, AO of the first manic/mixed episode, presence of psychosis, presence of incongruent and congruent psychosis, rapid cycling.

General age of onset of BP-I. In the combined sample (**Table 1**) the general AO was significantly predicted by SCZ3-PRS-CS (P=3.33E-04**)** and by all eight PTs (C+T method) (FDR-corrected P-values from P=0.0273 for PT_5E-08 to P=2x10^{-7} for PT_0.05). (**Table_S7**). In the RO sample the general AO was significantly predicted by SCZ3-PRS-CS (P = 3.36E-05), while in the UK sample just a trend was visible (P=0.0904). In both national samples the regression coefficients were negative indicating that a higher SCZ3-SNP loading was associated with a younger AO in BP-I patients.

Age of onset of the first depressive episode. Both in the combined RO-UK sample (P=1.96E-04) and in the separate national samples the age of onset of the first depressive episode was significantly predicted by the SCZ3-PRS-CS (P=1.96E-04 for RO; P=3.24E-02 for UK) **(Table 1)** Similarly, all eight SCZ3-PTs computed through the C+T method predicted the AO of depression with significant P-values and with negative regression coefficients in the RO-UK sample (**Table_S8**) and in the national samples (data not shown) indicating a negative effect of SCZ3-PRS on AO of depression.

The ***AO-depression was younger in psychotic*** patients than in non-psychotic patients in the RO-UK sample (AO-depression in psychotics X=25.90, sd=10.38, AO-depression in non-psychotics =27.04, sd=11.44; t=2.89; P=0.004), as well the general AO of BP-I (AO in psychotics M=25.05; sd=9.54; AO in non-psychotics M=26.35; sd =11.16 (t= 2.49; P=0.013).

Age of onset of the first manic episode was significantly predicted by the SCZ3-PRS-CS both in the combined sample (P=4.82E-05) and in the national samples **(**P=4.02E-04 for RO**;** P=9.68E-03 for UK) (**Table 1**). The C+T method did not predict the AO of the first manic episode either in the national or in the combined samples (data not shown).

In our samples there was a significant difference in AO of depression between female and male cases. (RO sample mean AO-depression: males M=30.27 years (s.d.=10.60); females M=27.11 years (s.d.=9.77; t=3.364, df=1/497, P=0.00082; UK sample: males M=26.45 (s.d.=11.33; females M=24.26 (s.d.=10.41; t=2.95, df =1/ 884, P=0.003); RO-UK sample AO-depression; males M=27.76, sd=11.23; females M=25.31, sd=10.27, df=1/1381; t=4.15, P=1.7x10-05). Therefore, we performed linear regressions with the proband sex as covariate in the combined sample. The covariate sex strongly influenced the effect of SCZ3-PRS on the AO-depression (FDR-P=3.00E-06), but the significance of SCZ3-PRS for the AO-depression was preserved in both methods (data not shown).

Presence of lifetime psychosis and incongruent psychosis. Similar to other BP samples (Aminoff et al., 2022) the prevalence of psychosis (congruent and incongruent) reached 71% in the RO-UK BP-I sample.

Both PRS computation methods (CS and C+T) yielded highly significant P-values for the prediction of psychosis irrespective of type and for the mood incongruent psychosis in the combined sample (**Table 1** and **Tables_S9-S10)** and the national samples (data not shown).

We found a relatively new result indicating a ***negative correlation between the number of depressive episodes and psychosis.*** This finding was confirmed by a multivariate logistic regression (B=-0.407; sd=0.143, Wald=8.113; OR=0.666, 95%CI=0.503-0.881, P=0.004) including six clinical variables, regularised regressions, RF **(Table 4**), The same negative correlation was valid for the incongruent psychosis, but not for congruent psychosis (**Table 4).**

The ***mood congruent psychosis*** was significantly predicted only by the SCZ3-PRS-CS in the combined sample (**Table 1),** but not by the C+T method.

Rapid cycling BP-I. Both the SCZ3-PRS-CS method (**Table 1)** and the C+T SCZ3-PRS method with five PTs (**Table_S11**) significantly predicted the rapid cycling trait. But the ORs were below 1 and the regression coefficients were negative suggesting that rapid cycling and SCZ3-PRS loading have an inverse relationship.

Family history for major psychoses (schizophrenia, schizoaffective disorders, bipolar disorder, unipolar major depression) was nominally predicted by three SCZ3-PTs **(Table_S12)** indicating that only specific SNPs and genes are involved in familial inheritance.

3.3. *Predictive performance of SCZ3-PRS in combination with clinical traits of BP-I*

After investigating the predictive power of SCZ3-PRS-CS for phenotypic traits of BP-I in univariate regressions (only the SCZ3-PRS-CS regressed against each outcome) we investigated the predictive power of SCZ3-PRS-CS in combinations with clinical variables (family history of major psychoses in first and second degree relatives, number of depressive and manic episodes, AO-mania, rapid cycling, irritable mania, total number of episodes, AO-depression, general AO) for certain BP-I traits in the RO-UK sample with the random forest method that controls the collinearity between predictor variables (**Table 2)**. BP-I cases were randomly allocated to either training, validation or testing sets. To determine the predictive performance, i.e., classification by the cross-validated RF model of the binary outcomes, the ROC and its AUC, sensitivity, specificity, and accuracy were used. Additionally, the **PPV** and the **F1** score were reported. Accuracy for the cross-validated RF regression of the continuous outcomes was assessed with $\mathbf{R^2}$ and **RMSE. Table 2** shows that both the accuracy and AUC-values for binary subphenotypes (psychosis and its subtypes) and R^2 and RMSE for continuous subphenotypes indicate a moderate predictive performance of SCZ3-PRS-CS and clinical variables. The best predictions were for psychosis, incongruent psychosis (AUC close to 0.8) and AO-depression, consistently across methods.

3.4. *Models of best prediction of phenotypic traits of BP-I by SCZ3-PRS alone and in combination with clinical data.*

As shown by Lewis and Vassos (2022) PRS alone do not provide 100% accuracy of phenotype prediction. Other factors such as family history of major psychiatric disorders and age-of-onset may influence the phenotypic development too.

Therefore, we explored prediction models for each BP-I trait comparing clinical variables or SCZ3-PRS-CS with models combining both clinical variables and PRS. Clinical variables included in the prediction models were: family history of major psychoses, total number of episodes, number of manic episodes, number of depressive episodes, irritable mania, rapid cycling, general AO, AO-depression, AO-mania. Each one of these variables was excluded from the predictors when it became the outcome.

For all investigated BP-I traits the best predictions were provided by models including combinations of SCZ3-PRS-CS and clinical variables followed by models containing only clinical variables. The worst prediction indicators appeared in models including only SCZ3-PRS-CS (**Table 3**). However, there were significant differences in the metrics between the clinical and clinical plus SCZ3-PRS models, tested in pairwise Bonferroni-corrected one-sample t-tests, except for congruent psychosis and AO-mania, for which only trends appeared. The best predictions were for psychosis and incongruent psychosis, AO-depression.

Because in all analyses, psychosis was the best predicted BP-I subphenotype, we tried to estimate the ***importance of the above mentioned variables*** for predicting psychosis and its subtypes examining the effect of two cross-validation methods: regularised regression elastic net in the **"cv.glmnet**" (cross-validation general linear model) and conditional random forest (RF) (**cforest**) in R **(Table 4).**

In RF the predictor variables with a higher score (MDA=mean decrease accuracy) represent more accuracy loss when excluding the variable, therefore indicating the variable is more important for the classification of the outcome. The observed permuted, cross-validated P-value indicates (with a $P < 0.05$) variable importance, rejecting the null distribution. **Table 4** shows that the importance of different variables for prediction changes according to the used method. While family history, SCZ3-PRS-CS, number of mania episodes, rapid cycling, irritable mania, general AO have the highest and equal importance in the elastic net model for the prediction of psychosis, in the RF model the importance of all these variables diminishes, while remaining significant. For ***mood-incongruent psychosis*** the most important predictors in both models were family history of major psychoses and irritable mania; SCZ3-PRS-CS and general AO of BP-I remain significant at a **higher** P-value.

3.5. *Pathway analysis of psychosis in BP-I*

Twenty-two pathways (**Table 5)** had a competitive P-value of ≤ 0.05, defined as showing enrichment. All significantly enriched pathways contained at least one gene identified in previous or most recent GWAS of SCZ (Trubetskoy et al., 2022) or BP (Mullins et al., 2021) **(Table 6).**

The highest associated pathways were ***ZNF318*** (R^2=0.951, FDR-P=0.003), ***Apoptosis*** (R^2=0.958; FDR-P=0.003), and ***Mitochondrion*** (R^2=0.754; FDR-P=0.037). ***ZNF318*** (zinc finger protein 318) was identified in the most recent BP GWAS (Mullins et al., 2021).

Other pathways associated with psychosis in our samples were pathways relevant to brain function, including synaptic transmission involving both ion channels and dendrites (***regulation_of_dendritic_spine_development; regulation_of_membrane_repolarization; regulation_of_dopamine_receptor_signaling)***, to the autonomous nervous system ***(abnormality_of_the_autonomic_nervous_system),*** to the immune system (***regulation_of_immune_system_process).***

4. DISCUSSION

A strength of our study is the phenotypic homogeneity of strictly diagnosed BP-I probands and the direct investigation of the controls with a psychiatric interview, which is not always the case of large-scale GWAS samples and the highlight of some new associations.

Our study is among the first checking the predictive validity of the SCZ3-PRS for BP-I clinical traits in phenotypically homogeneous clinical samples. There is only one published study using SCZ3-PRS for prediction of the clinical course of the disease in psychotic patients (mainly schizophrenia) (Landi et al., 2021) but not for predicting those clinical traits we investigated.

In our sample the most significant associations of the SCZ3-PRS were with psychosis, incongruent psychosis, and AO-depression in both PRS computation methods. We also evidenced a negative correlation between the number of depressive episodes and psychosis.

Our results confirm findings of previous studies that used the SCZ-SNP-set 2014 (Ripke et al., 2014) and SCZ3-SNP-set (Trubetskoy et al., 2022) regarding the effect of SCZ-SNPs on psychosis in BP-I (Leonenko et al., 2018; Richards et al., 2022; Allardyce et al., 2023) and on AO in BP-I (Ruderfer et al., 2018; Kalman et al., 2021). Moreover, a higher burden of SCZ3-PRS was

associated not only with younger general AO of BP-I, but also with decreased AO of first depressive episode and of first manic episode. A relationship between SCZ3-PRS and AO-depression was reported by Harder et al. (2022) for the AO of unipolar major depression in the UK biobank.

We investigated the predictive power of SCZ3-PRS-CS in combinations with other nine clinical variables in RF models that control for the collinearity between predictor variables. According to AUC and accuracy values for dichotomous traits and R^2 and RMSE for continuous traits the predictive power of SCZ3-PRS was more modest than in simple linear /logistic regressions, but the best prediction was for incongruent psychosis and AO-depression. Moreover, the RF models that compared the predictions based on only SCZ3-PRS, on SCZ3-PRS plus clinical variables, and on only clinical variables showed that the worst prediction was provided by the SCZ3-PRS and that the accuracy of the prediction based on only clinical variables was not far from that based on both SCZ33-PRS and clinical variables. This finding is in line with the observation of Landi et al., (2021) that SCZ3-PRS "did not improve individualized outcome prediction relative to information from a routine psychiatric assessment" and the observation was valid for different ancestries and ascertainment strategies of SCZ patients.

Lewis and Vassos (2022) and Pedersen et al. (2022) suggested that clinical variables like psychiatric family history and age of onset improve the predictions based on PRS for clinical purposes. This was evident inTables 3-4, in the ranking of clinical variables in the prediction of psychosis, as well as in the comparison of the RF model based on only SCZ3-PRS with the model including SCZ3-PRS plus clinical variables.

Our both PRS methods significantly confirmed the trend observed in the sample of Ruderfer et al. (2018) that there is an ***inverse relationship between SCZ-SNP loading and BP rapid cycling.***

To our knowledge the ***negative correlation between the number of depressive episodes and psychosis*** found in our samples both in regularised regressions and (RF) is a novel finding supported by a meta-analysis of 54 studies of psychotic symptoms in BP (Aminoff et al., 2022) showing that psychosis is four times more frequent in manic/mixed episodes than in depressive episodes of BP-I**.**

Contrary to Allardyce et al., (2023) who found no effect of SCZ3-PRS on mania in BP, we found a significant influence of SCZ3-PRS on the AO of mania. As a difference to our BP-I sample, their sample contained 28.8% BP-II cases who, by definition, have hypomanic episodes that may be sometimes triggered by antidepressant medication.

Enrichment analysis revealed several pathways previously related to schizophrenia and psychosis. The enriched pathways are relevant to brain function, including synaptic transmission involving both ion channels and dendrites, and brain development.

The pathways that explained the highest variance of psychosis were: ***ZNF318***, ***Apoptosis***, ***Mitochondrion***.

Neuroimaging studies showed progressive loss of cortical gray matter in first-episode psychosis (Jarskog et al., 2005**),** therefore a role for apoptosis mechanisms producing cell or localised synaptic/dendritic loss in psychosis is plausible. Defects in the structure of dendrites of pyramidal neurons may also have direct effects leading to the loss of cortical volume (***regulation_of_dendritic_spine_development***) (Garey et al, 1998**)** .

Mitochondrial dysfunction **(*Mitochondrion***) was linked to alterations in dopamine signaling, glutamatergic dysfunction and oxidative stress in schizophrenia (Ramaker et al, 2017; Whitehurst and Howes, 2022) and in BP (Andreazza et al., 2010**).**

Both the "***Mitochondrion***" and "***ZNF318"*** pathways contain the CREB3L4-gene. CREB3L4 is a subtype of the CREB1-gene, expression of which is downregulated in brain tissue of SCZ, BP, MDD patients compared with healthy controls (Xiao et al., 2018).

The ***chr1p21*-pathway** with the microRNA encoding gene ***MIR137HG*** that regulates signaling pathways for neural development is implicated in schizophrenia risk (Thomas et al., 2017).

NCOA2 (R^2=0.542, FDR-P=0.027) was one of 9 genes differentially expressed in the dorsolateral prefrontal cortex (DLPFC) in patients with BP (Nurnberger et al, 2014).

Regulation of dopamine was implicated in psychosis by the ***regulation_of_dopamine_ receptor_signaling_pathway***. Excessive dopaminergic modulation of striatal function has long been hypothesized to mediate psychosis and antipsychotic drugs target dopaminergic innervation in the striatum (Benjamin et al., 2022).

There is also evidence to involve the immune system in the pathogenesis of psychosis (***regulation_of_immune_system_process***). Increased risk of adulthood psychosis has been linked to high concentrations of proinflammatory cytokines in childhood (Perry et al., 2021). In a GWAS of response of BP patients to lithium treatment (International Consortium on Lithium Genetics, 2018) genes related to the immune system (HLA antigen complex and inflammatory cytokines) were associated with the treatment response and the same genes in the HLA region were also associated with risk for BP (Mullins et al., 2021) and SCZ (Trubetskoy et al., 2022).

In the "***negative_regulation_of_immune_system_process***" and the *"**ZNF318***" pathways appears the ***MAD1L1*-**gene and in the ***NOA2*-**pathway appears the ***NT5C2*-**gene that were associated with BP and SCZ in several GWAS; they were also associated with the AO of BP-I in the RO sample (Grigoroiu-Serbanescu et al., 2015). The immune system PRS pathway, further implicated the gene ***FURIN***, recently associated with BP (Mullins et al., 2021); it was linked with decreased neurite outgrowth (Deans et al., 2023; Schrode et al., 2019).

Altered function of the autonomic nervous system involving heart rate was previously documented in SCZ and psychosis (Kocsis et al., 2020; Montaquilla et al., 2015) and genes present in this pathway were associated with cardiac β-adrenergic signaling and cardiac hypertrophy signaling in BP (Nurnberger et al., 2014).

Several genes (***CACNA1C, GABBR1 , GABBR2, SLC6A9; NT5C2***) in pathways linked to psychosis in our BP-I sample are also involved in the epigenomic differential methylation of DNA in SCZ and psychosis (Hannon et al., 2021; Hannon et al., 2016). DNA methylation was also linked to the AO of SCZ (Srivastava et al., 2022) and BP-I (Grigoroiu-Serbanescu et al., 2012).

Comparison of two methods for calculating the SCZ3 individual-level PRS. A comparison of the variance explained (liability Nagelkerke R^2) by the two methods indicated a marked increase in phenotypic variance explained by PRS-CS compared to the C+T method (**Table_S13**). For some investigated BP-I traits (AO-mania, mood-congruent psychosis) the two PRS computation methods gave different results. On the other hand, the C+T method showed that PTs stringently associated SCZ3-SNPs (e.g. PT-5E-08; PT-1E-07) offer significant predictions for incongruent psychosis and AO-depression supporting their clinical validity.

Study limitations. A clinical limitation of the combined sample is a certain degree of heterogeneity of the BP-I severity: only the RO sample and a part of the UK sample included hospitalized BP-I patients. The hospitalization is an indicator of illness severity that may influence the presence of psychotic symptoms (Aminoff et al., 2022), number of episodes, and AO. Furthermore, as presented in Table S1, a part of the UK cases had no information about several subphenotypes of BP-I.

5. Conclusion

Our study is among the first investigating the predictive value of the SCZ3-PRS and shows that they have a modest clinical value for predicting some phenotypic traits of BP-I in machine learning models and that their best performance is in combination with clinical variables, although the analysed traits of BP-I were significantly predicted in univariate regressions. We showed that the risk for psychosis, mood-incongruency and early onset age, particularly onset of depression are associated with high genetic burden for SCZ3-PRS and these variables together with the SCZ3-PRS might be used in the clinical counselling for BP-I treatment since previous studies using SCZ-PRS derived from SCZ-GWAS (2014) showed that a high burden of SCZ-PRS is associated with poor

response to antipsychotic and lithium treatment (Zhang et al., 2019; International Consortium on Lithium Genetics**,** 2018; Schubert et al., 2021).

We also showed that the results may be influenced by the method of computing the PRS and by the statistical models used.

Author contribution

MGS and AM planed the project, recruited the participants, and assured the funding. TvV, TB, MGS conducted statistical analyses. CCD, AIN, SH, JT, AJF, MMN performed genomic analyses. NB participated to the recruitment of UK participants. MGS and TvV wrote the paper. All authors reviewed and edited the manuscript.

Acknowledgements

The UK participants' genetic and clinical data were collected with support from the Neuroscience Research Charitable Trust, the Central London NHS (National Health Service) Blood Transfusion Service, the Camden and Islington NHS Foundation Trust, a research lectureship from the Priory Hospitals and the National Institutes for Health Research (NIHR) Mental Health Research Network (MHRN).

Funding:

The Romanian sample was funded by UEFISCDI, Bucharest, Romania, grant number 203/2021 (PN-III-P4-ID-PCE-2020-2269) to Maria Grigoroiu-Serbanescu.

Recruitment of a proportion of the bipolar disorder and control subjects and for genotyping of the bipolar disorder and control samples were supported by the Stanley Center for Psychiatric Research at the Broad Institute. Participant recruitment was also funded by the Medical Research Council (MRC grant code G1000708). Andrew McQuillin and Nick Bass were supported by the UCLH NIHR BRC.

Declaration of Competing Interest

None of the authors have any conflict of interest to declare.

Declaration of Generative AI and AI assisted technologies in the writing process.

No AI and AI assisted technologies were used in the writing process or statistical analyses.

References

Allardyce, J., Leonenko, G., Hamshere, M., Pardinas, A.F., Forty, L., Knott, S., Gordon- Smith, K., Porteous, D.J., Haywood, C., Di Florio, A., Jones, L., McIntosh, A.M., Owen, M.J., Holmans, P., Walters, J.T.R., Craddock, N., Jones, I., O'Donovan, M.C., Escott-Price, V., 2018. Association between schizophrenia-related polygenic liability and the occurrence and level of mood-incongruent psychotic symptoms in bipolar disorder. JAMA Psychiatry 75, 28–35. https://doi.org/10.1001/ jamapsychiatry.2017.3485.

Allardyce, J., Cardno, AG., Gordon-Smith, K., Jones L, Di Florio, A., Walters, J.T.R., Holmans, PA., Craddock, NJ., Jones, I., Owen, M.J., Escott-Price, V., O'Donovan, M.C., 2023. Specificity of polygenic signatures across symptom dimensions in bipolar disorder: an analysis of UK bipolar disorder research network data. Lancet Psychiatry, 10 (8), 623-631. doi: 10.1016/S2215-0366(23)00186-4.

American Psychiatric Association. 1994. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). American Psychiatric Press, Washington, DC.

Aminoff, S.R., Onyeka, I.N., Ødegaard, M., Simonsen, C., Lagerberg, T.V., Andreassen, O.A., Romm, K.L., I.,Melle, I., 2022. Lifetime and point prevalence of psychotic symptoms in adults with bipolar disorders: a systematic review and meta-analysis. Psychol. Medicine, 52 (13), 2413-2425. https://doi.org/ 10.1017/S003329172200201X**.**

Andreazza, AC., Shao, L., Wang, J.F., Young, L.T., 2010. Mitochondrial complex I activity and oxidative damage to mitochondrial proteins in the prefrontal cortex of patients with bipolar disorder. Arch. Gen. Psychiatry. 2010; 67(4):360-368. https://doi.org//10.1001/archgenpsychiatry.2010.22

Benjamin, K.J.M., Chen, Q., Jaffe, A.E., Stolz, J.M., Collado-Torres, L., Huuki-Myers, L.A., Burke, E.E., Arora, R.,, Feltrin, A.S., Rocha Barbosa, A., Radulescu**,** E., Pergola, G., Heon Shin, J., Ulrich, W.S., Deep-Soboslay, A., Tao , R., BrainSeq Consortium, Hyde, T.M., Kleinman, J.E., Erwin, J.A., Weinberger, D.R., Paquol, A.C.M., 2022. Analysis of the caudate nucleus transcriptome in individuals with schizophrenia highlights effects of antipsychotics and new risk genes. Nature Neurosci. 25, 1559–1568**.** https://doi.org/10.1038/s41593-022-01182-7.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerfull approach to multiple testing. J. Royal Statistical Society. Series B, 57, 289-300.

Calafato, M.S., Thygesen, J.H., Ranlund, S., Zartaloudi, E., Cahn, W., Crespo-Facorro, B., Díez-Revuelta, A., Di Forti, M., Genetic Risk and Outcome of Psychosis (GROUP) consortium, Hall, M-H, Iyegbe, C., Jablensky, A., Kahn, R., Kalaydjieva, L., Kravariti, E., Lin, K., McDonald, C., McIntosh, A.M., McQuillin, A., Psychosis Endophenotypes International Consortium (PEIC), Picchioni, M., Rujescu, D., Shaikh, M., Toulopoulou, T., Van Os, J.,

Vassos, E., Walshe, M., Powell, J., Lewis, C.M., Murray, R.M., Bramon, E., Wellcome Trust Case Control Consortium (WTCCC2), 2018. Use of schizophrenia and bipolar disorder polygenic risk scores to identify psychotic disorders**.** Brit. J. Psychiatry 213, 535–541. https://doi.org/ 10.1192/bjp.2018.89.

Chang, C.C., Chow, C.C., Tellier, L., C.A.M., Vattikuti Sh, Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer dataset. 2015. GigaScience, 4, (1), s13742-015-0047-8, https://doi.org/10.1186/s13742-015-0047-8.

Chen, B-Y., Lin, J-J., Lu, M-K., Tan, H-P., Jang, F-L., Lin, S-H., 2021. Neurodevelopment regulators miR-137 and miR-34 family as biomarkers for early and adult onset schizophrenia. NPJ Schizophrenia 7, 35; https://doi.org/10.1038/s41537-021-00164-1.

Choi, S.W., O'Reill, P.F., 2019. PRSice-2 Polygenic Risk Score software for biobank-scale data. 2019. GigaScience, 8, 1–6. https://doi.org/10.1093/gigascience/giz082.

Choi, S.W., García-González, J., Ruan, Y., Wu, H.M., Porras, C., Johnson, J; Bipolar Disorder Working group of the Psychiatric Genomics Consortium; Hoggart, C.J., O'Reilly PF., 2023. PRSet: Pathway-based polygenic risk score analyses and software. PLoS Genetics, 19 (2), e10.1371/journal.pgen.1010624. eCollection 2023**.**

Cross-Disorder Group of PGC, 2019. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. Cell, 179, 1469-1482.e11. https://doi.org/10.1016/j.cell.2019.11.020.

Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M., Schlessinger, D., Stambolian, D., Loh P-R, Iacono, W.G., Swaroop, A., Scott, L.J., Cucca, F., Kronenberg, F., Boehnke, M., Abecasis, G.R., Fuchsberger, C., 2016. Next-generation genotype imputation service and methods**.** Nat. Genet. 48, 1284-1287. https://doi.org/ 10.1038/ng.3656.

Deans, P.J.M, Seah, C., Johson, J., …K.J.,Brennand. 2023. Non-additive effects of schizophrenia risk genes reflect convergent downstream function**.** medRxiv. Preprint. https://doi.org/10.1101/2023.03.20.23287497

De Leeuw, C.A., Mooij, J.M., Heskes, T., Posthuma, D., 2015. MAGMA: Generalized gene-set analysis of GWAS data. PLoS Computational Biology, 11(4): e1004219. https://doi.org/10.1371/journal.pcbi.1004219.

De Hond, A.A.H., Steyerberg, E.W., van Calster, B., 2022. Interpreting area under the receiver operating characteristic curve. The Lancet - Digital Health 4 (12), e853-e854. doi:https://doi.org/10.1016/S2589-7500 (22)00188-1.

Endicott, J., Spitzer R.L., 1978. A diagnostic interview: the schedule for affective disorders and schizophrenia. Arch Gen Psychiatry. 35 (7):837-844. doi: 10.1001/archpsyc.1978.01770310043002

Euesden, J.C.M., O'Reilly, P.F., 2015. PRSice: Polygenic Risk Score software . Bioinformatics 31 (9):1466-468. https://doi.org/10.1093/bioinformatics/btu848

Finucane, H.K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P-R., Anttila, V., Xu, H., Zang, C., Farh, K., Ripke, S., Day, F.R., ReproGen Consortium; Schizophrenia Working Group of the Psychiatric Genomics Consortium; RACI Consortium; Purcell, S., Stahl, E., Lindstrom, S., Perry, J.R.B., Okada, Y., Raychaudhuri, S., Mark, J., Daly, M.J., Patterson , N., Neale, B.M., Price, A .L., 2015. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat. Genet. 47, 1228–1235. https://doi.org/10.1038/ng.3404.

Garey, L.J., Ong, W.Y., Patel, T.S., Kanani, M., Davis, A., Mortimer, A.M., Barnes, T.R., Hirsch, S.R., 1998. Reduced dendritic spine density on cerebral cortical pyramidal neurons in schizophrenia. J. Neurol. Neurosurg. Psychiatry 65 (4):446-453. https://doi.org/10.1136/jnnp.65.4.446.

Ge, T., Chen, Ch-Y., Ni, Y., Feng, A.Y-Ch., Smoller, J.W., 2019. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat. Commun. 10, 1776. https://doi.org /10.1038/s41467-019-09718-5.

Grigoroiu-Serbanescu, M., Diaconu, C.C., Heilmann-Heimbach , S., Neagu, A.I., Becker. T., 2015. Association of age-of-onset groups with GWAS significant schizophrenia and bipolar disorder loci in Romanian bipolar I patients. Psychiatry Research 230 (3): 964-967. https://doi.org/ 10.1016/j.psychres.2015.11.008.

Grigoroiu-Serbanescu, M., Wickramaratne, P.J., Mihailescu, R., Prelipceanu, D., Sima, D., Codreanu, M., Grimberg, M., Elston, R.C., 2012. Paternal age effect on age of onset in bipolar I disorder is mediated by sex and family history. Am. J. Med. Genet. B Neuropsychiatr. Genet. 59B (5):567-579. doi: 10.1002/ajmg.b.32063.

Grigoroiu-Serbanescu, M., Giaroli, G., Thygesen, J.H., Shenyan, O., Bigdeli, T.B., Bass, N.J., Diaconu, C.C., Neagu, A.I., Forstner, A.J., Degenhardt, F., Herms, S., Nöthen, M.M., McQuillin, A., 2020. Predictive power of the ADHD GWAS 2019 polygenic risk scores in independent samples of bipolar patients with childhood ADHD**.** J. Affect. Disord. 265, 651-659. doi: 10.1016/j.jad.2019.11.109.

Grotzinger, A.D., Mallard, T.T., Akingbuwa, W.A., Ip H.F., Adams, M.J., Lewis, C.M., McIntosh AM, Grove, J., Dalsgaard, S., Lesch, K.P., Strom, N., Meier, S.M., Mattheisen, M., Børglum, A.D., Mors, O., Breen, G.; iPSYCH; Tourette Syndrome and Obsessive Compulsive Disorder Working Group of the Psychiatric Genetics Consortium; Bipolar Disorder Working Group of the Psychiatric Genetics Consortium; Major Depressive Disorder Working Group of the Psychiatric Genetics Consortium; Schizophrenia Working Group of the Psychiatric Genetics Consortium; Lee, P.H., Kendler, K.S., Smoller, J.W., Tucker-Drob, E.M., Nivard, M.G., 2022. Genetic architecture of 11 major psychiatric disorders at biobehavioral, functional genomic and molecular genetic levels of analysis. Nat. Genet. 54 (5), 548–559. https://doi.org/10.1038/s41588-022-01057-4

Hannon, E., Dempster, E.L., Mansell, G., Burrage, J., Bass, N., Bohlken, M.M., Corvin, A., Curtis, C.J., Dempster, D., Di Forti M., Dinan, T.G., Donohoe, G., Gaughran, F., Gil, M., Gillespie, A., Gunasinghe, C., Hulshoff, H.E., Hultman, C.M., Johansson, V., Kahn, R.S., Kaprio, J., Kenis, G., Kowalec, K., MacCabe, J., McDonald, C., McQuillin, A., Morris, D.W., Murphy, K.C., Mustard, C.J., Nenadic, J., O'Donovan, M.C., Quattrone, D. , Richards, A.L., Rutten, B.P.F., St Clair, D., Therman, S., Toulopoulou, T., Van Os, J., Waddington, J.L., Wellcome Trust Case Control Consortium (WTCCC), CRESTAR consortium, Sullivan, P., Vassos, E. , Breen, G., Collier, D.A., Murray, R.M., Schalkwyk, L., Mill, J., 2021. DNA methylation meta-analysis reveals cellular alterations in psychosis and markers of treatment-resistant schizophrenia. Elife 26, 10:e58430. https://doi.org/10.7554/eLife.58430.

Hannon, E., Dempster, E., Viana, J., Burrage, J., Smith, A.R., Macdonald, R.,St Clair, D., Mustard, C., Breen, G., Therman, S., Kaprio, J., Toulopoulou, T., Hulshoff Pol, H.E., Bohlken, M.M., Kahn, R.S., Nenadic, I., Hultman, C.M., Murray, R.M., Collier, D.A., Bass, N., et al. 2016. An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization of genetic associations and differential DNA methylation. Genome Biol. 17, 176. doi: https://doi.org/10.1186/s13059-016-1041-x.

Harder, A., Nguyen, T-D., Pasman, J.A., Mosing, M.A., Hägg, S., Lu, Y., 2022. Genetics of age-at-onset in major depression. Transl. Psychiatry 12 (1):124. https://doi.org/10.1038/s41398-022-01888-z.

https://github.com/getian107/PRScs.

Humpston, C. S., Bebbington, P., Marwaha, S., 2021. Bipolar disorder: Prevalence, help-seeking and use of mental health care in England. Findings from the 2014

Adult Psychiatric Morbidity Survey. J Affect. Disord. 262, 426-433. doi: 10.1016/j.jad.2020.12.151.

International Consortium on Lithium Genetics (ConLi+Gen), 2018. Association of polygenic score for schizophrenia and HLA antigen and inflammation genes with response to lithium in bipolar affective disorder: a genome-wide association study. JAMA Psychiatry 75 (1), 65-74. https://doi.org/10.1001/jamapsychiatry.2017.3433.

Jarskog, L.F., Glantz, L.A., Gilmore, J.H., Lieberman, J.A., Apoptotic mechanisms in the pathophysiology of schizophrenia**.**. 2005. Prog. Neuropsychopharmacol. Biol. Psychiatry 29 (5):846-58. https://doi.org / 10.1016/j.pnpbp.2005.03.010.

Joo, J.W., Hormozdiari, F., Han, B., Eskin, E., 2016. Multiple testing correction in linear mixed models. Genome Biol. **17, 62**. https://doi.org/10.1186/s13059-016-0903-6

Kalman, J.L., Loohuis, O.L.M., Vreeker, A., McQuillin, A., Stahl, E.A., Ruderfer, D., Grigoroiu-Serbanescu, M., Panagiotaropoulou, G., Ripke, S., Bigdeli, T.B., Stein, F., Meller, T., Meinert, S., Pelin, H., Streit, F., Papiol, S., Adams, M.J., Adolfsson, R., Adorjan, K., Agartz, I., Aminoff, SR., Anderson-Schmidt, H., Andreassen, O.A., Ardau, R., Aubry, Saghatolislam, F., Knowles, JA., Kranz, T.M., Lagerberg, TV., Landen, M., Lawson, W.B., Leboyer, M., …….. Di Florio, A., Smoller, J., Biernacka, J.M., McMahon, F., Alda, M., Müller-Myhsok, B., Koutsouleris, N., Falkai, P., Freimer N, Andlauer ,T., Schulze, T., Ophoff, R. A. Characterisation of age and polarity at onset in bipolar disorder**.** 2021**.** Brit. J. Psychiatry 219 (6), 659-669 . https://doi.org /10.1192/bjp.2021.102.

Kalman, J,L,, Papiol, S., Forstner, A.J., Heilbronner, U., Degenhardt, F., …., Fullerton, J.M., Gade, K., Gard, S., Garnham, J.S., Goes, F., Grigoroiu-Serbanescu, M., Gryaznova, A1., Hake, M1., Hauser, J30., Herms, S., …… Nöthen, M., Rietschel, M., Schulze, T.G., 2019. Investigating polygenic burden in age at disease onset in bipolar disorder: Findings from an international multicentric study. Bipolar Disord. 21, 68-75. https://doi.org /10.1111/bdi.12659.

Kalman, J.L., Loohuis, O., Vreeker, L.M., McQuillin, A., Stahl, E.A., Ruderfer, D., Grigoroiu-Serbanescu, M., Panagiotaropoulou, G., Ripke, S., Bigdeli, T.B., Stein, F., Meller, T., Meinert, S., Pelin, H., Streit, F., Papiol, S; Adams, M.J., Adolfsson, R., Adorjan, K., Agartz, I., Aminoff, S.R., Anderson-Schmidt, H., Andreassen, O.A., Ardau, R., Aubry, …… Di Florio, A., Smoller, J., Biernacka, J.M., McMahon F, Alda M, Müller-Myhsok, B., Koutsouleris, N., Falkai, P., Freimer, N., Andlauer, T., Schulze, T., Ophoff, R. A**.** 2021. Characterisation of age and polarity at onset in bipolar disorder**.** Brit. J. Psychiatry 219 (6), 659-669. https://doi.org/ 10.1192/bjp.2021.102.

Kocsis, A., Gajwani, R., Gross, J., Gumley, A.J., Lawrie, S.M., Schwannauer, M., Schultze-Lutter, F., Grent-'t-Jong, T., Uhlhaas, P.J., 2020. Altered autonomic function in individuals at Clinical high risk for psychosis. Front. Psychiatry 11:580503 . .https://doi.org/10.3389/fpsyt.2020.580503eCollection 2020.

Kuhn, M., (2008). Building predictive models in R using the caret package. J. Stat. Software 28 (5), 1–26. https://doi.org/10.18637/jss.v028.i05

Landi, I., Kaji, D.A., Cotter, L., Van Vleck, T., Belbin, G., Preuss, M., Loos, R.J.F., Kenny, E., Glicksberg, B.S., Beckmann, N.D., O'Reilly, P., Schad,t E.E., Achtyes, E.D., Buckley, P.F., Lehrer, D., Malaspina, D.P., McCarroll, S.A., Rapaport, M.H., Fanous, A.H., Pato, M.T., Pato, C.N., Bigdeli, T.B., Nadkarni, N.G., Charney, A.W., 2021. Prognostic value of polygenic risk scores for adults with psychosis**.** Nat. Med. 27 (9),1576-1581. https://doi.org/10.1038/s41591-021-01475-7

LeDell, E., Petersen, M., van der Laan, M., (2022). Cross-validated area under the ROC curve confidence intervals version 1.1.4. http://CRAN.R-project.org/package=cvAUC

Lee, S.H., Goddard, M.E., Wray, N.R., Visscher, P.M., 2012. A better coefficient of determination for genetic profile analysis. Genet. Epidemiol. 36 (3):214-24. https://doi.org/ 10.1002/gepi.21614. PMID: 22714935.

Leeuw, C.A., de Mooij, J.M., Heskes, T., Posthuma, D., 2015. MAGMA: Generalized gene-set analysis of GWAS data. PLOS Comput. Biol. 11: e1004219. .https://doi.org/ 10.1371/journal.pcbi.1004219 .

Leonenko, G., Di Florio, A., Allardyce, J., Forty, L., Knott, S., Jones, L., Gordon-Smith, K., Owen, M.J., Jones, I., Walters, J., Craddock, N., O'Donovan, M.C., Escott-Price, V., 2018. A data-driven investigation of relationships between bipolar psychotic symptoms and schizophrenia genome-wide significant genetic loci. Am. J. Med. Genet. B Neuropsychiatr. Genet. 177 (4):468-475. https://doi.org/ 10.1002/ajmg.b.32635.

Lewis, C. M., Vassos, E., 2022**.** Polygenic Scores in Psychiatry: On the road from discovery to implementation. Am. J. Psychiatry, 179 (11): 800-806. . https://doi.org/

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., Mesirov, JP., 2011. Molecular signatures database (MSigDB) 3.0. Bioinformatics 27 (12):1739-40. .https://doi.org /10.1093/bioinformatics/btr260.

Loh, P. R., Danecek, P., Palamara, P.F., Fuchsberger, C., Reshef, Y.A., Finucane, H.K., et al , 2016. Reference-based phasing using the Haplotype Reference Consortium panel. Nat. Genet. 48, 1443–1448.

McGuffin, P., Farmer, A., Harvey, I.,1991. A polydiagnostic application of operational criteria in studies of psychotic illness. Development and reliability of the OPCRIT system. Arch. Gen. Psychiatry 48 (8): 764-770. https://doi.org/10.1001/archpsyc.1991.01810320088015

McMahon, F.J., 2014. Prediction of treatment outcomes in psychiatry—where do we stand ? Dialogues Clin. Neurosci, 16 (4), 455–464. doi: 10.31887/DCNS.2014.16.4/fmcmahon

Merikangas, K.R., Jin, R., He, J-P., Kessler, R.C., Lee, S., Sampson, N.A., Viana, M. C., Andrade, L.H., Hu,C., Karam, E.G., Ladea, M., Medina-Mora, M.E., Ono,Y., Posada- Villa, J., Sagar, R., Wells, E., Zarkov, Z., 2011. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. Arch. Gen. Psychiatry, 68 (3), 241-251. https://doi.org/ 10.1001/archgenpsychiatry.2011.12.

Montaquila, J.M., Trachik, B.J., Bedwell, J.S.,2015. Heart rate variability and vagal tone in schizophrenia: a review. J. Psychiatr. Res. 69: 57–66. https://doi.org/ 10.1016/j.jpsychires.2015.07.025

Mullins, N., Forstner, A.J., O'Connell, K.S., Coombes, B., Coleman, J.R., Qiao, Z., Thomas, D. Als, Bigdeli,T.B., Børte, S., Bryois, J., Charney, A.W,, Drange, O. K., Gandal, M. J., ,… …. …B., Etain,P., Ferentinos, M., Frye, J.M., Fullerton, M., Gawlik, E.S., Gershon, F., Goes, M. J., Green, Grigoroiu-Serbanescu, M., J., Hauser, F., Henskens, J., Hillert, K., Sue Hong, D.M., Hougaard, C.M., Hultman, K., Hveem, N., Iwata, A.V., Jablensky, Jones, I., …….Werge, T., Wray, N., Zwart, J-A., Biernacka, J. A., Nurnberger, J. I., Cichon, S., Edenberg, H., Stahl, E.A.A., McQuillin, A., Di Florio, A., Ophoff, R., Andreassen, O.A., Genome-wide association study of over 40,000 bipolar disorder cases provides novel biological insights. Nat. Genet. 2021, 53 (6) , 817 -829 **.** https://doi.org/10.1038/s41588-021-00857-4

Nurnberger, J.I., Blehar, M.C., Kaufmann, C.A., York-Cooler, C., Simpson, S.G., Harkavy-Friedman, J., et al, 1994. Diagnostic interview for genetic studies (DIGS). Rationale, unique features, and training. Arch. Gen. Psychiatry 51, 849-859.

Nurnberger, Jr. J.I., Koller, D.L., Jung, J., Edenberg, H.J., Foroud, T., Guella, I., Vawter, M.P., Kelsoe, J.R., Psychiatric Genomics Consortium Bipolar Group, 2014. Identification of pathways for bipolar disorder: a meta-analysis. JAMA Psychiatry, 71(6), 657-64. doi: 10.1001/jamapsychiatry.2014.176.

Örd, T., Lönnberg,T., Nurminen,V., Ravindran,A., Niskanen,H., Kiema, M., Õunap, K., Maria,M., Moreau, P.R., Mishra, P.P., Palani, S., Virta, J., Liljenbäck, H., Aavik, E., Roivainen, A., Ylä-Herttuala, S., Laakkonen, J.P., Lehtimäki, T., Kaikkonen, M.U., 2023. Dissecting the polygenic basis of atherosclerosis via disease-associated cell state signatures. Am. J. Hum. Genet. 110 (5), 722-740. doi: 10.1016/j.ajhg.2023.03.013.

Pain, O., Dudbridge, F., Ronald, A.,2018. Are your covariates under control? How normalization can re-introduce covariate effects. Eur. J. Hum. Genet. 26, 1194–1201. https://doi.org/10.1038/s41431-018-0159-6

Pain, O., Glanville,K.P., Hagenaars, S.P., Selzam, S., Fürtjes, A.E., Gaspar,H.A., Coleman,J.R.I., Rimfeld, K., Breen, G., Plomin, R., Folkersen, L., Lewis, C.M., 2021. Evaluation of polygenic prediction methodology within a reference-standardized framework. PLoS Genet. 4, 17 (5) e1009021. doi: 10.1371/journal.pgen.1009021. eCollection

Pain, O., Gillett, A.C., Austin, J.C., Folkersen, L., Lewis,C.M., 2022. A tool for translating polygenic scores onto the absolute scale using summary statistics. Eur. J. Hum. Genet. 30 (3), 339-348. doi: 10.1038/s41431-021-01028-z.

Pedersen, E.M., Agerbo, E., Plana-Ripoll, O., Grove, J., Dreier, J.W., Musliner, K.L., Bækvad-Hansen, M., Athanasiadis,G., 11Schork, A., Bybjerg-Grauholm, J., Hougaard, D.M., Werge,T., Nordentoft, M., Mors, O., Dalsgaard, S., Christensen, J., Børglum, A.D., Mortensen, P.B., McGrath,J.J., Prive, F., Vilhja´lmsson., B.J., 2022. Accounting for age of onset and family history improves power in genome-wide association studies. Am. J. Hum. Genet. 109, 1–16.

Perry, B.I., Zammit, S., Jones, P.B., Khandaker, G.M., 2021. Childhood inflammatory markers and risks for psychosis and depression at age 24: Examination of temporality and specificity of association in a population-based prospective birth cohort. Schizophr. Res. 230, 69-76. doi: 10.1016/j.schres.2021.02.008.

Pistis, G., Vázquez-Bourgon, J., Fournier, M., et al., 2022. Gene set enrichment analysis of pathophysiological pathways highlights oxidative stress in psychosis. Mol. Psychiatry 27, 5135–5143**.** https://doi.org/10.1038/s41380-022-01779-1

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., Reich, D., 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38 (8), 904-909. doi: 10.1038/ng1847 .

Ramaker, R.C., Bowling, K.M., Lasseigne, B.N., Hagenauer, M.H., Hardigan, A.A., Davis, N.S., Gertz, J., Cartagena, P.M., Walsh, D.M., Vawter, M.P., Jones, E.G., Schatzberg,A.F., Barchas, J.D., Watson, S.J., Bunney, B.G., Akil, H., Bunney,W.E., Li, J.Z., Cooper, S.J., Myers, R.M., 2017. Post-mortem molecular profiling of three psychiatric disorders **.** Genome Medicine 9 (1) 72. doi: 10.1186/s13073-017-0458-5.

R Core Team, 2020 . R**:** A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from http://www.r-project.org/index.html

Richards, A.L., Cardno, A., Harold, G., Craddock, N.J., Di Florio,A., Jones,L., Gordon-Smith,K., Jones,I., Sellers, R., Ph.D., Walters, J.T.R., Holmans, P.A., Owen, M J.,

O'Donovan, M.C., 2022. Genetic liabilities differentiating bipolar disorder, schizophrenia, and major depressive disorder, and phenotypic heterogeneity in bipolar disorder. JAMA Psychiatry, 79 (10):1032-1039. doi:10.1001/jamapsychiatry.2022.2594

Ripke S et al., Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014. Biological insights from 108 schizophrenia-associated genetic loci. Nature 511 (7510), 421-427.

Ruderfer DM et al, Bipolar Disorder & Schizophrenia Working Group of PGC, 2018. Genomic dissection of bipolar disorder and schizophrenia, including 28 subphenotypes. Cell 173, 1705-1715.e16. https://doi/ 10.1016/j.cell.2018.05.046.

Schrode, N., Ho., S.M., Yamamuro, K., Dobbyn, A., Huckins, L., Matos, M.R., Cheng, E., Deans, P.J.M., Flaherty, E., Barretto, N., Topol, A., Alganem, K., Abadali, S., Gregory, J., Hoelzli, E., Phatnani, H., Singh, V., Girish, D., Aronow, B., Mccullumsmith, R., Hoffman, G.E., Stahl, E.A., Morishita, H., Sklar, P., Brennand, K.J., 2019. Synergistic effects of common schizophrenia risk variants. Nat. Genet. 51 (10), 1475-1485. doi: 10.1038/s41588-019-0497-5

Schubert, K.O., Thalamuthu, A., Amare, A.T., Frank, J., Streit, F., Adli, M., Akula, N., Akiyama, K., Ardau, R., Arias, B., Aubry, J.M., Backlund, L., Bhattacharjee, A.K., Bellivier, F., Benabarre, A., Bengesser, S., Biernacka, J.M., Birner, A., ……..Chillotti, C., Cichon, S., Clark, S.R., Cruceanu, C., Czerski, P.M., Dalkner, N., Dayer, A., Degenhardt, F., DelZompo, M., DePaulo, J.R., Étain, B., Falkai, P., Forstner, A.J., Frisen, L., Frye, M.A., Fullerton, J.M., Gard, S., Garnham, J.S., Goes, F.S., Grigoroiu-Serbanescu, M., Grof, P., …… Reininghaus ,E., Rouleau, G.A., Rybakowski, J.K., Schalling, M., Schofield, P.R., Schweizer, B.W., Severino, G., Shekhtman, T., Shilling, P.D., Shimoda, K., Simhandl, C., Slaney, C.M., Squassina, A., Stamm, T., Stopkova, P., Tekola-Ayele, F., Tortorella, A., Turecki, G., Veeh, J., Vieta, E., Witt, S,H., Roberts, G., Zandi, P., Alda, M., Bauer, M., McMahon, F.J., Mitchell, P.B., Schulze, T.G., Rietschel, M., Baune, B.T., 2021. Combining schizophrenia and depression polygenic risk scores improves the genetic prediction of lithium response in bipolar disorder patients. Transl. Psychiatry, 11(1), e606. https://doi.org/ 10.1038/s41398-021-01702-2.

Smoller, J.W., Finn, C.T., Family, twin, and adoption studies of bipolar disorder, 2003. Am. J. Med. Genet. Part C Semin. Med. Genet. 123C, 48–58.

Stahl, E.A., Breen, G., Forstner, A.J., McQuillin, A., Ripke, S., Trubetskoy, V., et al., 2019. Genome-wide association study identifies 30 loci associated with bipolar disorder. Nat. Genet. 51, 793–803. https://doi.org/ 10.1038/s41588-019-0397-8.

Srivastava, A., Chaudhary, Z., Qian J., Al Chalabi, N., Burhan A.M., Fischer, C.E., Gerretsen, P., Kolla, N.J., Graff A., Remington, G., De Luca, V., 2022. Genome-wide methylation analysis of early-onset schizophrenia. Psychiatr. Genet. 32(6), 214-220. doi: 10.1097/YPG.0000000000000317.

Takahashi, N., Nishimura, T., Harada, T.,, A., Choi, D., Iwabuchi, T., Kuwabara, H., Takagai, S., Nomura, Y., Newcorn, J.H., Takei, N., Tsuchiya, K.J., 2020. Polygenic risk score analysis revealed shared genetic background in attention deficit hyperactivity disorder and narcolepsy. Transl. Psychiatry 10 (1), 284. doi: 10.1038/s41398-020-00971-7.

The 1000 Genomes Project Consortium. 2015**.** A global reference for human genetic variation. Nature 526, 68–74. https://doi.org/10.1038/nature15393.

Thomas, K.T., Anderson, B.R., Shah, N., Zimmer, S.E., Hawkins, D., Valdez, A.N., Gu, Q., Bassell, G.J., 2017. Inhibition of the Schizophrenia-Associated MicroRNA miR-137 Disrupts Nrg1α Neurodevelopmental Signal Transduction. Cell Rep. 20 (1), 1-12. https://doi.org/ 10.1016/j.celrep.2017.06.038.

Trubetskoy, V., Pardiñas, A.F., Qi, T., Panagiotaropoulou, G., Awasthi, S, Bigdeli,T.B., Bryois,J., Chen, Ch-Y., Dennison, Ch.A., Hall, L.S.S., Walters, J.T.R., O'Donovan, M.C., Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2022. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature 604 (7906), 502-508. doi: 10.1038/s41586-022-04434-5.

Whitehurst, T., Howes, O., 2022. The role of mitochondria in the pathophysiology of schizophrenia: A critical review of the evidence focusing on mitochondrial complex one. Neurosci. Biobehav Rev. 32 :449-464. https://doi.org /10.1016/j.neubiorev.2021.11.047.

Xiao, X., Zhang, C., Grigoroiu-Serbanescu, M., Wang, L., Li, L., Zhou, D., Yuan, T-F., Wang, C., Chang, H., Wu, Y., Li, Y., Wu, D-D., Yao, Y-G., Li, M., 2018 . The cAMP responsive element-binding (CREB)-1 gene increases risk of major psychiatric disorders. Mol. Psychiatry 23 (9):1957-1967. https://doi.org/ 10.1038/mp.2017.243.

Zhang, J-P., Robinson, D., Yu, J, Gallego, J., Fleischhacker,W., Kahn, R.S., Crespo-Facorro, B, Vazquez-Bourgon, J, Kane, JM, Malhotra, A.K., Lencz ,T. 2019. Schizophrenia polygenic risk score as a predictor of antipsychotic efficacy in first-episode psychosis. Am. J. Psychiatry 176: 21–28; doi: 10.1176/appi.ajp.2018.17121363.

Table 1. PRS-SCZ3 linear and logistic regressions predicting BP-I case status and BP-I subphenotypes in the RO, UK, and RO-UK samples

PRS-CS-auto SCZ3 score (Predictor)				Effect Size		Variance %	AUC	P	FDR-P
Sample	**Outcome**	**Beta**	**SE**		**95%CI**				
	BP-I vs Ctrl			**OR**		**R^2N [1]**			
RO-UK	BPI	0.533	0.033	1.705	1.71; 1.82	4.261	0.643	2.84E-64	4.83E-63
RO	BPI	0.638	0.068	1.893	1.89; 2.17	5.901	0.669	3.55E-23	1.21E-22
UK	BPI	0.514	0.038	1.672	1.67; 1.80	3.963	0.639	1.43E-44	1.22E-43
	Psychosis								
RO-UK	Psychosis	0.463	0.035	1.589	1.48; 1.70	3.264	0.624	1.53E-42	2.14E-41
RO	Psychosis	0.576	0.068	1.778	1.56; 2.03	4.915	0.654	2.89E-19	1.35E-18
UK	Psychosis	0.443	0.042	1.557	1.43; 1.69	2.981	0.619	4.45E-27	3.12E-26
				OR		**R^2N [1]**			
RO-UK	Incongruent	0.464	0.058	1.591	1.42; 1.79	3.285	0.626	3.30E-16	1.16E-15
RO-UK	Congruent	0.381	0.057	1.464	1.31; 1.64	2.243	0.603	1.03E-11	2.88E-11
RO-UK	Rapid cycling	-0.451	0.382	0.637	0.61; 0.69	2.014	0.583	2.13E-09	2.43E-08
				t-test		**Adj. R2**			
RO-UK	AO general	-0.944	0.256	-3.68	-1.45; -0.44	0.728	-	2.38E-04	3.33E-04
RO	AO general	-1.88	0.431	-4.38	-2.73; -1.04	0.935	-	1.44E-05	3.36E-05
UK	AO general	-0.535	0.315	-1.69	-1.15; 0.08	0.472	-	0.0904	9.04E-02
				t-test		**Adj. R2**			
RO-UK	AO depress	-1.2	0.313	-3.84	-1.81; 0.59	0.98	-	1.26E-04	1.96E-04
RO	AO depress	-1.92	0.495	-3.88	-2.89; 0.95	1.197	-	1.17E-04	1.96E-04
UK	AO depress	-0.86	0.396	-2.17	-1.64; 0.08	0.526	-	3.01E-02	3.24E-02
				t-test		**Adj. R2**			
RO-UK	AO mania	-1.34	0.316	-4.24	-1.96; 0.72	0.731	-	2.41E-05	4.82E-05
RO	AO mania	-1.83	0.504	-3.62	-2.82; 0.84	0.815	-	3.16E-04	4.02E-04
UK	AO mania	-1.07	0.403	-2.65	-1.86; 0.28	0.363	-	8.30E-03	9.68E-03

Abbreviation. AO general – BP-I age of onset ; AO depres – Age of onset of depression; Incongruent – mood incongruent psychosis; [1] Nagelkerke pseudo R-squared on the liability scale

Table 2. PRS-SCZ3 prediction of BPI -traits with 10-fold cross-validation out-of-sample (random forest classification)

Outcome	Sample	Accuracy	Accuracy. CI (95%)	PPV	F1	AUC	AUC/ CI (95%)
Psychosis	RO/UK	0.765	0.737-0.799	0.777	0.852	0.785	0.74-0.82
Incongruent	RO/UK	0.805	0.769-0.836	0.819	0.884	0.787	0.75-0.83
Congruent	RO/UK	0.724	0.685-0.760	0.74	0.803	0.761	0.72-0.80
		RMSE	**R2**	**MAE**	**RMSE-sd**	**R2 sd**	**MAE-sd**
AO BPI	RO/UK	5.591	0.733	3.716	0.647	0.057	0.357
AO Depression	RO/UK	3.884	0.874	2.105	0.62	0.048	0.265
AO Mania	RO/UK	5.082	0.774	2.801	0.698	0.063	0.314
All models used conditional inference **random forest** to reduce risk of overfitting in models with correlated predictors.							

Abbreviations: Adj.R2 = adjusted R2; RMSE = Root Mean Squared Error; MAE = Mean Absolute Error, PPV=positive prediction value; AUC = Area Under the Curve

Table 3. Random forest 10- fold cross-validated predictions using clinical only, clinical and genetic, and genetic only (SCZ3-PRS-CS) predictors in the RO-UK sample

Model (RO-UK sample)	Accuracy	Accuracy CI (95%)	AUC	AUC 95% CI	Adj. P[1]
Psychosis					
Clinical + genetic	0.765	0.737-0.799	0.785	0.744-0.824	-
Clinical	0.719	0.704-0.794	0.761	0.722-0.817	7.32E-03
Genetic (SCZ3-PRS)	0.711	0.685-0.733	0.625	0.593- 0.657	
Incongruent psychosis					
Clinical + genetic	0.805	0.769-0.836	0.787	0.746-0.829	-
Clinical	0.786	0.752- 0.779	0.753	0.644-0.724	4.59E-04
Genetic (SCZ3-PRS)	0.738	0.714-0.761	0.606	0.575-0.638	
Congruent psychosis					
Clinical + genetic	0.724	0.685-0.760	0.761	0.719-0.803	-
Clinical	0.696	0.65-0.706	0.693	0.643-0.714	0.096
Genetic (SCZ3-PRS)	0.685	0.659-0.710	0.601	0.568-0.634	
	RMSE	**Adj. R^2**			
AO BP-I					
Clinical + genetic	5.591	0.733	-	-	-
Clinical	6.279	0.714	-	-	< 2.2E-16
Genetic (SCZ3-PRS)	10.651	0.016			
AO depression					
Clinical + genetic	3.884	0.874	-	-	-
Clinical	5.037	0.768	-	-	8.00E-03
Genetic (SCZ3-PRS)	9.874	0.017			
AO mania					
Clinical + genetic	5.082	0.774	-	-	-
Clinical	6.143	0.681	-	-	0.078
Genetic (SCZ3-PRS)	10.766	0.015			

[1] Bonferroni corrected P for pairwise comparisons (*"caret"* R-package). The adj. P-value represents a significant pairwise difference between the 3 model AUCs (SCZ3-PRS alone, clinical, and clinical + SCZ3-PRS); the p-value is adjusted for 3 repeats (models) of 10-fold validations, i.e. for 30 results.
Abbreviations: Adj.R2 = adjusted R2; RMSE = Root Mean Squared Error; MAE = Mean Absolute Error, AUC = Area Under the Curve.

Table 4. Comparison between variable importance in Random Forest and regularised regression for prediction of psychosis and psychosis subtypes

Regularised regression								Conditional Random forest (RF)		
		'cv.glmnet' (Elastic Net model) with bootstrap SE, CI, and P-value						**'cforest' (conditional RF)** with conditional permutation importance		
Outcome	**Predictor**	**[1]OR**	**SE**	**CI.low**	**CI.high**	**P.value**	**Sig.**	**[2]MDA**	**P**	**Sig.**
Psychosis	FH major psychoses	1.769	0.138	1.48	2.076	< 0.001	***	0.0216	0.011	*
Psychosis	SCZ3-PRS	0.194	0.051	0.1	0.293	< 0.001	***	0.003	0.009	**
Psychosis	Nr depres episodes	-0.132	0.066	-0.244	-0.011	0.028	*	0.002	0.002	**
Psychosis	Nr epis mania	0.29	0.109	0.232	0.627	< 0.001	***	0.002	0.031	*
Psychosis	Age of onset mania	-0.014	0.013	-0.039	0.011	0.204	n.s.	0.0002	0.017	*
Psychosis	Rapid cycling	-0.746	0.325	-1.293	-0.02	< 0.001	***	0.002	0.027	*
Psychosis	Irritable mania	2.479	0.225	2.053	2.954	< 0.001	***	0.002	0.052	
Psychosis	Nr epis total	0.108	0.026	0.015	0.12	0.012	*	0.001	0.023	*
Psychosis	Age of onset depression	-0.037	0.015	-0.065	-0.009	0.016	*	0.0001	0.024	*
Psychosis	Age onset BPI	-0.069	0.017	-0.108	-0.041	< 0.001	***	0.0003	0.029	*
Mood-incongruent	SCZ3 PRS	0.261	0.047	0.211	0.404	< 0.001	***	0.008	0.012	*
Mood-incongruent	FH major psychoses	2.325	0.142	2.158	2.744	< 0.001	***	0.006	0.009	**
Mood-incongruent	Irritable mania	1.739	0.226	1.295	2.165	< 0.001	***	0.001	0.003	**
Mood-incongruent	Age of onset depression	-0.041	0.017	-0.067	-0.002	0.02	*	0.0005	0.021	*
Mood-incongruent	Age of onset mania	-0.009	0.013	-0.032	0.018	0.456		0.001	0.021	*
Mood-incongruent	Nr epis mania	0.197	0.039	0.13	0.268	0.004	**	0.001	0.028	*
Mood-incongruent	Age onset BPI	-0.086	0.013	-0.117	-0.064	< 0.001	***	0.0001	0.027	*
Mood-incongruent	Nr epis depres	-0.067	0.111	-0.313	0.04	0.4		0.003	0.033	*
Mood-incongruent	Rapid cycling	-0.083	0.197	-0.412	0.39	0.776		0.0005	0.009	**
Mood-incongruent	Nr epis total	0.109	0.019	0.077	0.158	< 0.001	***	0.0003	0.04	*
Mood-congruent	FH major psychoses	1.85	0.119	1.575	2.033	< 0.001	***	0.01	0.013	*
Mood-congruent	SCZ3 PRS	0.173	0.051	0.068	0.265	< 0.001	***	0.005	0.009	**
Mood-congruent	Nr epis depres	0.27	0.053	0.192	0.367	0.004	**	0.002	0.03	*
Mood-congruent	Nr epis mania	0.384	0.042	0.307	0.478	< 0.001	***	0.001	0.024	*
Mood-congruent	Age of onset mania	-0.075	0.014	-0.12	-0.056	< 0.001	***	0.001	0.026	*
Mood-congruent	Rapid cycling	-0.277	0.305	-0.427	-0.122	< 0.001	***	0.0002	0.004	**
Mood-congruent	Irritable mania	1.68	0.186	1.358	2.121	< 0.001	***	0.002	0.022	*
Mood-congruent	Nr epis total	0.2	0.022	0.158	0.249	< 0.001	***	0.0004	0.034	*
Mood-congruent	Age of onset depression	-0.11	0.019	-0.178	-0.106	< 0.001	***	0.001	0.027	*
Mood-congruent	Age onset BPI	-0.118	0.016	-0.171	-0.108	< 0.001	***	0.001	0.043	*
[1]Log OR per SD; [2]MDA, Mean decreased accuracy; FH = family history										

Table 5. PRSet SCZ3 Individual level pathway analysis in the RO-UK sample

Pathway PRS for psychosis	Association with psychosis		Pathway enrichment for psychosis	
	R2*(%)	FDR P-value[a]	Nr. SNPs	P-value[b]
MITOCHONDRION	0.754	0.037	4899	0.010
ZNF318	0.951	0.003	4333	0.002
REGULATION_OF_IMMUNE_SYSTEM_PROCESS	0.339	0.038	1378	0.043
NCOA2	0.542	0.027	1609	0.011
MIR202_3P	0.54	0.043	1367	0.009
MIR3125	0.7	0.014	1432	0.004
MIR6859_5P	0.634	0.015	1186	0.006
MIR4782_5P	0.665	0.014	1041	0.004
MIR5706	0.665	0.014	1041	0.004
MIR4763_3P	0.659	0.014	1012	0.005
ABNORMALITY_OF_THE_AUTONOMIC_NERVOUS_SYSTEM	0.68	0.014	1022	0.004
MIR10395_3P	0.63	0.015	535	0.005
REGULATION_OF_DENDRITIC_SPINE_DEVELOPMENT	0.607	0.019	458	0.005
MIR197	0.658	0.014	433	0.004
REGULATION_OF_MEMBRANE_REPOLARIZATION	0.548	0.026	336	0.007
APOPTOSIS	0.958	0.003	539	0.000
chr1p21	0.649	0.014	359	0.004
REGULATION_OF_DOPAMINE_RECEPTOR_SIGNALING_PATHWAY	0.317	0.014	136	0.018
MIR625_3P	0.48	0.038	200	0.009
MIR3681_5P	0.636	0.015	223	0.004
MIR6849_5P	0.683	0.014	230	0.003
MIR4669	0.445	0.036	28	0.009

SNP= single nucleotide polymorphism.

*Pathways presented are the weighted R^2 , i.e. R2 devided by the number of SNPs in the pathway.

[a] P-values for association after FDR multiple testing correction, significance was set at $p < 0.05$.

[b] P-values indicating enrichment were corrected for 10,000 permutations, significance was set at $p < 0.05$.

Table 6. Significant GWAS genes associated with psychosis included in the 22 enriched pathways

22 Pathway PRS for psychosis	Genes in pathway identified in GWAS				
1.MITOCHONDRION	BRD8 TRIM31 (**BP**) CKB SFXN2 CLU CLIC1 (**BP**) GLYCTK (**BP**) SLC9B2 LETM2 **CREB3L4 (SCZ)**	METTL15 MARK2 (**BP**) ALAS1 (**BP**) FEN1 (**BP**) FHIT MLXIP (**SCZ**) FOXO3 HARS2 (**BPI**) **GABBR1 (BP)** HSPA9	HSPD1 HSPE1 IRF3 YJEFN3 (**BPI**) FADS1 (**BP**) MAPT MSRA (**BP**) NDUFA2 (**BPI**) **NRGN** CISD2 (**SCZ**)	PCCB (**SCZ**) TMX2 MRPS33 (**BP**) PLEC (**BP**) POLG MIEF1 NDFIP2 (**BP; SCZ**) DARS2 (**SCZ**) NDUFAF7 AMBRA1	DNAJC11 MAPK3 (**BP; SCZ**) STARD7 (**BP**) ELAC2 SDHAF1 NDRG4
2.ZNF318	CAMKK2 PTK2B CHRNA2 MATN4 (**BPI**) SPECC1 **SYNE1 (BP)** TMTC1 (**SCZ**) DOCK2 (**BP**) OASL SLC39A1	GPM6A (**SCZ**) TDRD9 CDC25C (**BPI**) NEGR1 (**SCZ**) TBL1XR1 ZNF664 HARBI1 NMB (**BPI**) CTNND1 TCTN1	MEF2C-A MYO19 (**BPI**) MAD1L1 (**BPI;SCZ**) DNAJC11 PLEKHO1 (**BP**) SDCCAG8 SMG6 IGSF9B WDR76 (**SCZ**) FOXP1	DARS2 (**SCZ**) ENOX (**SCZ**) KDM3B (**BP**I) TBC1D5 UBE2D2 RC3H1 SEC11A (**BP**I) RERE (**SCZ**) **CREB3L4 (SCZ)** GATAD2B	DOC2A MSI2 (**SCZ**) SPPL3 ZEB2 ATG13 (**SCZ**) GRIN2A (**BP; SCZ**) DLGAP2 (**SCZ**)
3.REGULATION_OF_IMMUNE_ SYSTEM_PROCESS	PLK2 RC3H1 DRD2 PLCL2 SCRIB HSPA9 MDK	**FURIN (BP)** YTHDF2 **MAD1L1** (**BP; SCZ)** CUL4A DGKZ CD4			
4. NCOA2 TARGET GENES	ITIH1 (**BPI**) NT5C2 (**SCZ; BPI**) PACSIN2 (**BP**) OGFOD2 HSPA9	IPO13 ALOX5AP DYNC1LI2 RBMS3 ABCB9	RBKS RELA AGPAT1 SNHG3	MEF2C MSANTD2	

5.MIR202_3P	MEF2C PLEKHO1 (**BP**) PTPRD (**SCZ**) BCL7A ZNF823 (**SCZ**) SHANK2 (**SCZ**) MOB4 (**BPI**)	MSI2 (**SCZ**) SH3RF3 (**BPI**) RD3L HSPE1-MOB4 TSPAN2			
6.MIR3125	ELAVL4 ZNF365 (**BP**) RC3H1 CUL4A (**BP**)	ANKRD45 ARL3 SUFU TRIM8			
7.MIR6859_5P	ELAVL4 RC3H1 NEBL ARL3	SUFU TRIM8			
8.MIR4782_5P	ADD3 **(BP)** ALAS1 **(BP)** RPS6KA 2 **(BP)** TCTN1 **(BPI)**	CALN1 (**SCZ**) SUMO2 (**BP**) DNMT3A DYNC1LI2			
9.MIR5706	ADD3 (**BP**) ALAS 1 (**BP**) RPS6KA2 **(BP**) TCTN1	CALN 1 (**SCZ**) SUMO2 (**BP**) DNMT3A DYNC1LI2			
10.MIR4763_3P	SLC6A9 DEF8 ETF1 NGEF KIF21B	TAF12 SUFU GATAD2B KLF6 (**SCZ**) STAG1	MLXIP (**SCZ**) SMARCD1 MEF2C MARK2 (**BP**)		
11.ABNORMALITY_OF_THE_ AUTONOMIC_NERVOUS_SYSTEM	MAPT (**SCZ**) **CACNA1C (BPI;** **SCZ**) ZEB2 CHRNA3 FGFR1 TUBB3	GIGYF2 ARL3 TCTN1 (**BPI**) SNCA **GABBR2 (SCZ)**	FANCI FANCA FANCL CISD2 (**SCZ**) SUFU		
12.MIR10395_3P	ADD 3 (**BP**) ATXN7 **(BPI)** CHRNA5 KLF6 **(SCZ)**	MOB4 (**SCZ**) HSPE1-MOB4	CADM2 (**BPI)** ACE (**SCZ)**		

13.REGULATION_OF_DENDRITIC_ SPINE_DEVELOPMENT	NGEF MEF2C SHANK2 **(BPI)**				
14.MIR197	BCL7A CTNND1 GATAD2B	SPPL3 ETF1			
15.REGULATION_OF_MEMBRANE_ REPOLARIZATION	YWHAE (**BP**) AKAP6				
16.APOPTOSIS	BNIP3L (**SCZ**) CLU RELA	DPYD (**SCZ**) ETF1			
17.chr1p21	DPYD (**SCZ**) NFU1P2 RPL7P9	RN7SKP270 PTBP2 **MIR137HG**			
18.REGULATION_OF_DOPAMINE_ RECEPTOR_SIGNALING_PATHWAY	DRD2				
19.MIR625_3P	WDR76 (**SCZ)** PTPRD (**SCZ**) ALAS1 **(BP)**				
20.MIR3681_5P	CSDE1 FUT10				
21.MIR6849_5P	CSDE1				
22.MIR4669	CLIC1 **(BP)** GIGYF2				