Associations between polygenic liability for schizophrenia and level of psychosis and mood-incongruence in bipolar disorder

Judith Allardyce[1] MRCPsych PhD, Ganna Leonenko[1] PhD, Marian Hamshere[1] PhD, Antonio F. Pardiñas[1] PhD, Liz Forty[1] PhD, Sarah Knott[2] PhD, Katherine-Gordon Smith[2] PhD, David J. Porteus[3] PhD, Caroline Haywood[3] PhD, Arianna Di Florio[1] MD PhD, Lisa Jones[2] PhD, Andrew M. McIntosh[3] FRCPsych MD, Michael J. Owen[1] FRCPsych PhD, Peter Holmans[1] PhD, James T.R. Walters[1] MRCPsych PhD, Nicholas Craddock[1] FRCPsych PhD, Ian Jones[1] FRCPsych PhD, Michael C. O'Donovan[1] FRCPsych PhD, Valentina Escott-Price[1] PhD

[1]MRC Centre for Neuropsychiatric Genetics and Genomics, Institute of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, UK

[2] Department of Psychological Medicine, University of Worcester, Worcester, UK

[3] Medical Genetics Section, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK

Division of Psychiatry, University of Edinburgh, Edinburgh, UK

Correspondence to: Judith Allardyce

MRC Centre for Neuropsychiatric Genetics & Genomics, Cardiff University, Hadyn Ellis Building, Maindy Road, Cathays, Cardiff, CF24 4HQ

+44 (0)29 2068 8320

allardycej2@cardiff.ac.uk

Word count: 2973

Date of revision 19/09/2017

Key Points

Question: what is the relationship between schizophrenia related polygenic liability and the occurrence and level of mood-incongruence of psychotic symptoms in bipolar disorder (BD)?

Findings: in this case-control study including 4436 BD cases, 4976 schizophrenia cases and 9012 controls, there was an exposure-response gradient of polygenic risk: Schizophrenia > BD with prominent mood-incongruent psychotic features > BD with mood-congruent psychotic features > BD with no psychosis, all differential associations were statistically-significant.

Meaning: A gradient of genetic liability across schizophrenia and bipolar disorder, indexed by the occurrence of psychosis and level of mood-incongruence has been shown for the first time.

Abstract

Importance

Bipolar disorder (BD) overlaps schizophrenia in its clinical presentation and genetic liability. Alternative approaches to patient stratification beyond current diagnostic categories are needed to understand the underlying disease processes/mechanisms.

Objectives

To investigate the relationship between common-variant liability for schizophrenia, indexed by polygenic risk scores (PRS) and psychotic presentations of BD, using clinical descriptions which consider both occurrence and level of mood-incongruent psychotic features.

Design

Case-control design: using multinomial logistic regression, to estimate differential associations of PRS across categories of cases and controls.

Settings & Participants

4399 BDcases, mean [sd] age-at-interview 46[12] years, of which 2966 were woman (67%) from the BD Research Network (BDRN) were included in the final analyses, with data for 4976 schizophrenia cases and 9012 controls from the Type-1 diabetes genetics consortium and Generation Scotland included for comparison.

Exposure

Standardised PRS, calculated using alleles with an association p-value threshold < 0.05 in the second Psychiatric Genomics Consortium genome-wide association study of schizophrenia, adjusted for the first 10 population principal components and genotyping-platform.

Main outcome measure

Multinomial logit models estimated PRS associations with BD stratified by (1) Research Diagnostic Criteria (RDC) BD subtypes (2) Lifetime occurrence of psychosis.(3) Lifetime mood-incongruent psychotic features and (4) ordinal logistic regression examined PRS associations across levels of mood-incongruence. Ratings were derived from the Schedule for Clinical Assessment in Neuropsychiatry interview (SCAN) and the Bipolar Affective Disorder Dimension Scale (BADDS).

Results

Across clinical phenotypes, there was an exposure-response gradient with the strongest PRS association for schizophrenia (RR=1.94, (95% C.I. 1.86, 2.01)), then schizoaffective BD (RR=1.37, (95% C.I. 1.22, 1.54)), BD I (RR= 1.30, (95% C.I. 1.24, 1.36)) and BD II (RR=1.04, (95% C.I. 0.97, 1.11)). Within BD cases, there was an effect gradient, indexed by the nature of psychosis, with prominent mood-incongruent psychotic features having the strongest association (RR=1.46, (95% C.I. 1.36, 1.57)), followed by mood-congruent psychosis (RR= 1.24, (95% C.I. 1.17, 1.33)) and lastly, BD cases with no history of psychosis (RR=1.09, (95% C.I. 1.04, 1.15)).

Conclusion

We show for the first time a polygenic-risk gradient, across schizophrenia and bipolar disorder, indexed by the occurrence and level of mood-incongruent psychotic symptoms.

Introduction

Although classified as a discrete diagnostic category[1-3], bipolar disorder (BD) overlaps considerably with schizophrenia (SCZ) in both its clinical presentation [4-13] and genetic liability [14-22]. BD is a phenomenologically heterogeneous construct and within the diagnostic category, individuals may have quite different symptom profiles. It has been proposed, that this clinical heterogeneity indicates underlying aetiological heterogeneity and the degree of clinical similarity between BD and SCZ reflects, overlapping alleles which selectively influence specific, shared clinical characteristics, rather than the global risk for the disorders [23-25].

Delusions and hallucinations are common in BD [26,27] with around one third of all psychotic features judged to be mood-incongruent [28,29]. Mood-incongruent psychotic features, are associated with poorer prognosis, poor lithium-response and are qualitatively similar to the prototypic symptoms of SCZ [30-32], suggesting that BD with psychosis and particularly mood-incongruent psychotic features, may specify a subgroup/stratum with stronger aetiological links to SCZ. Stratified linkage and candidate-gene studies of BD associations with chromosomal regions and genes implicated in SCZ, show stronger effects in psychosis and mood-incongruent subsamples [33-36] providing some support for this causal heterogeneity hypothesis, however lack of consistency in earlier linkage and candidate-gene studies renders the overall support weak.

Recently, genome-wide association studies (GWAS) have found a substantial polygenic component to both BD and SCZ risk, with a large proportion of their genetic variance explained by common alleles, partially shared across the two disorders [20]. Polygenic-risk can be calculated for individuals, with a single summary measure: the polygenic

risk score (PRS), which allows us to examine the genetic basis of symptom domains, within and across the two disorders [37-39] with greater power than the historical linkage and candidate-gene approaches. PRS-SCZ differentiate BD from controls [20,40] and there are differential associations across subtypes with schizoaffective bipolar disorder (SABD) (intermediate subtype, characterised by admixture of SCZ and BD symptoms) having a relatively larger burden of SCZ risk, compared to other BD subtypes [15,41]. To date, lack of power in well phenotyped samples has hindered fine-scale examination of the relationship between SCZ polygenic-risk and psychotic symptoms in BD.

We aimed to examine the relationship between polygenic liability for SCZ and psychotic presentations of BD using PRS generated from the most powerful SCZ-GWAS discovery set available, currently [21]. Measures relevant to the occurrence and nature of psychotic symptoms were considered. We hypothesised BD with psychosis would be associated with higher polygenic-risk for SCZ and this association would be stronger when mood-incongruent psychotic features were present, given their phenotypic similarity to the psychotic symptoms of prototypic SCZ.

Methods

Sample Ascertainment

Bipolar Disorder sample

4436 cases of BD with deep phenotypic information, European ancestry, domicile in the UK, collected between 2000 - 2013 were available via the UK BD Research Network (BDRN) using recruitment methods reported previously [15,42,43]. The sample has 1399 cases not included in prior BDRN publications [15,41]. All participants were assessed using a consistent protocol which included the Schedule for Clinical Assessment in Neuropsychiatry interview (SCAN) [44] administered by trained research psychologists

and psychiatrists, with very good to excellent inter-rater reliability for all domains of psychopathology [45]. Using information from the SCAN and casenote review, the Operational Criteria Checklist (OPCRIT) [46] was completed. Research Diagnostic Criteria (RDC) [3]diagnoses, which differentiate individuals on the basis of the their pattern of mood and psychotic symptoms better [41] than either DSM [2] or ICD–10[1], were made using the consensus lifetime best-estimate method, informed by all available information[47].

Schizophrenia sample

To allow comparison of BD with SCZ, we included a subset (N=4976) of the CLOZUK sample, collected via the Zapronex® Treatment Access System as detailed in a previous report[48], All were prescribed clozapine for treatment resistant SCZ (TRS) and are independent of, and unrelated (pi-hat <0.2) to individuals in the discovery GWAS[21]. In principle, TRS may carry higher polygenic-risk burden, however PRS in CLOZUK are similar to the other SCZ samples used by the Psychiatric Genomics Consortium[21].

Control Samples

The controls came from two UK sources: the Type-1 diabetes genetics consortium (TIDGC) (n = 2,532) are unscreened controls, recruited through the 1958 birth-cohort [49] and the other is a subsample of the Generation Scotland (n = 6,480) study, screened for psychiatric disorders [50]. Controls are unrelated (pi-hat < 0.2) to individuals in the PGC-SCZ discovery set, and were matched ancestrally to our case datasets [48].

All samples have appropriate ethics approvals.

Genotyping, quality control (QC), phasing and imputation

Bipolar cases

Genotypic data for the BD cases were processed in 3 batches, each on a different platform. To mitigate against potential bias from batch effects[51], stringent QC was

154 performed on each platform separately prior to merging. Single nucleotide
155 polymorphisms (SNPs) were excluded if the call rate was < 98%, MAF was < 0.01 or
156 they deviated from HWE at $p < 1x10^{-6}$. Individuals were excluded if they had minimal or
157 excessive autosomal homozygosity (|F|>0.1), high pairwise relatedness (pi-hat > 0.2)
158 or mismatch between recorded and genotypic sex. Following QC, the data for each
159 platform were phased using SHAPEIT [52] and imputed with IMPUTE2 [53], using the 1000
160 Genomes reference panel (Phase3, 2014). Imputed data were converted into the most
161 probable genotypes (probability >0.9) and merged on shared SNPs. 4399 BD cases
162 remained after QC.

163 ### CLOZUK cases and Controls

164 The CLOZUK and control samples had been though strict QC separately, before being
165 phased and imputed simultaneously as part of a larger SCZ study [48].

166 ### Merging BD, CLOZUK and control imputed genotypic datasets

167 After excluding SNPs with stand ambiguity; BD, CLOZUK and control samples were
168 merged and the imputed markers underwent a second QC filter[51], excluding SNPs with;
169 missingness in >5% of individuals, (INFO) <0.8, MAF <0.01 or deviation from HWE at p
170 $< 1x10^{-6}$.

171 ## Principal Component Analysis

172 To adjust for potential confounding from population structure, we performed PCA using
173 PLINK v1.9, after LD pruning and frequency filtering the SNPs from the merged sample,
174 keeping the eigenvectors for the first 10 principal components (PCs) to use as
175 covariates in the association analysis.

Polygenic Risk Scores (PRS)

We generated PRS[20], using the 2014 PGC-SCZ meta-analysis as our discovery set[21] calculated for each individual, based on a set of alleles with association p-values < 0.05. This decision was informed by the PGC leave one-cohort-out PRS analyses, for all SNP selection p-value thresholds, which found the median and mode of the cut-off = 0.05. This represents the association that best optimises the balance of false and true risk alleles, at the current discovery sample size [21]. The most informative and independent markers were selected to minimise statistical noise where possible, using p-value informed clumping, at $r^2 < 0.2$ with 1MB windows and by excluding the extended MHC (Chr6: position 25-35MB) because of its complex LD structure .

Outcome measure of lifetime psychosis & mood incongruence

Subtypes of BD

RDC subtypes were used as categorical outcomes in case-control analyses. The RDC [3] and Diagnostic and Statistical Manual of Mental Disorders (DSM) [2], though not the ICD-10 Classification of Mental and Behavioural Disorder (ICD-10) [1], subdivides BD into bipolar I disorder (BD I) and bipolar II disorder (BD II) depending on the nature of the mood states; mania in (BP I) and hypomania in (BP II). All classification systems recognise SABD. Psychotic symptoms are most prominent in SABD, then BD I, and least prominent in BD II [54,55].

The Bipolar Affective Disorder Dimension Scale

Outcome measures were generated from The Bipolar Affective Disorder Scale (BADDS) Psychosis (P) and mood-incongruence (I) subscales, which provide an ordered (not necessarily linear) measure of lifetime symptom domain severity[56]. An inter-rater

reliability exercise for this sample demonstrates excellent interclass correlation: (P) 0.91 and (I) 0.89.

1) A binary categorical outcome measure for lifetime occurrence of psychosis defined as an unambiguous episode of positive and/or disorganised psychotic symptoms, generated by dichotomising the (P) domain scale at a score > 9 [56].

2) A binary categorical outcome measure for lifetime occurrence of predominant mood-incongruent psychotic features (high v low prominence of mood-incongruence), generated by dichotomising the (I) domain scale at a score >19.

3) An ordinal measure of mood-incongruent psychotic features which assesses the overall balance between mood-congruent and mood-incongruent psychosis across the lifetime, rated using all available information according to BDRN protocol (E supplement : Note 1)

Statistical Analysis

A multinomial logit model (MNLM) was used to estimate differential associations of standardised PRS, adjusted for the first 10 PCs and genotyping-platform, across categories of cases and controls. We report the estimated coefficients transformed to relative risk-ratios (RR), defined as the exponentiated regression coefficient. PRS association across levels of mood-incongruent psychotic features using ordinal logistic regression was also estimated. To examine whether SABD subtypes were driving observed PRS associations with mood-incongruent psychotic features, we did a sensitivity analysis excluding SABD cases. Post-estimation predicted probabilities were plotted to aid interpretation of the PRS associations across RDC subtypes of BD [57]. To correct for multiple comparisons of PRS associations across different phenotypic strata within each model, bootstrapped standard errors and 95% confidence intervals were

generated, as an approximation to exact permutation methods [58](supplementary E - Note 2). Possible family-wise type-1 error proliferation was controlled for using the Bonferroni Method, calculated by multiplying the bootstrapped p-values by four [59].

Post-hoc analyses used a MNLM case-control design to examine differential associations across composite phenotypic categories defined by subtype BDI and BD II stratified by psychosis status and a complementary logistic regression analyses comparing the effect of PRS on lifetime occurrence of psychosis, across BD I and BD II subtypes. To examine the distribution of RDC defined cases across levels of PRS, we converted PRS to deciles and generated a stacked bar-chart (SCZ (CLOZUK), SABD, BD I, BD II), by decile.

Analyses were performed using PLINK v1.9 [60] or STATA (*Stata Statistical Software: Release 14*. College Station, TX: Stata Corp, LP).

Results

Sample description, Genotyping and quality control

After merging BD, CLOZUK and control imputed-genotyped samples and further QC, 18,387 cases and controls (E-supplementary Table 1) with 3,451,354 SNPs with INFO score > 0.8 and MAF >1% were available for analysis. Within the BD sample 52% (N = 2296) of cases endorsed lifetime occurrence of definite psychosis, with <1% missingness in this variable (N=25). Of the BD cases with definite psychosis, 43% (N= 981) were classed as having high lifetime mood-incongruent psychotic features. There was a 9% (N=214) missingness rate for the mood-incongruence variable, within the BD cases with psychosis.

Case Control PRS associations

As expected (Table 1 Section A), PRS discriminated CLOZUK from controls. PRS in those with a diagnosis of SABD or BD I, but not BD II, were significantly higher than controls.

PRS associations within cases

PRS discriminated SCZ from all BD subtypes (Table 2). Within BD, PRS discriminates BD II from both BD I and SABD (Figure 1). The percentage of CLOZUK cases increased monotonically with increasing decile PRS, while the percentage of bipolar subtypes decreased (Figure 2).

PRS associations with psychotic BD

Compared to controls, the PRS were higher in BD, regardless of whether there was a history of psychosis (Table 1, Section B, Figure 2). However, PRS were significantly higher in BD with psychosis, compared to BD without psychosis (Table 1, Section B, figure 3). Within BD cases, PRS discriminated those with and without psychosis (RR=1.25, 95% bootstrapped adjusted p-value < 001, C.I. (1.16, 1.33)).

Post hoc analyses showed the association between PRS and psychosis was present in BD I (OR = 1.21, 95% C.I. 1.10, 1.32) but were not statistically significant in BD II (OR = 0.98, 95% C.I. 0.80, 1.18). Composite subgroup defined as BD I with psychosis - had higher PRS compared to controls (RR = 1.38, 95% C.I. 1.31, 1.46) this association was significantly stronger than that of the composite BD I/no psychosis (RR= 1.16, 95% C.I. 1.08, 1.25). Within BD II, there was no differential association across subgroups defined by presence/absence of psychosis as compared to controls (supplementary-E: Table-1).

PRS associations with mood-incongruent psychotic features

Psychotic BD characterised by high mood-incongruence has a higher SCZ polygenic risk burden than controls, with a one standard-deviation increase in PRS increasing the RR of being in the high mood-incongruence category by 46% (RR= 1.46, bootstrapped, 95% C.I. 1.36, 1.57) (Figure 3, Table 1 Section C). Although the association was significantly weaker than for the high mood-incongruent group, schizophrenia risk-alleles were enriched in those with low mood-incongruence compared with controls (RR= 1.24, bootstrapped 95% C.I. (1.17, 1.33). Sensitivity analysis excluding the SABD group from analyses found comparable results (Table 1: Section D). Finally, a within-BD-case analysis, measuring mood-incongruence on an ordinal scale found the odds of having higher levels of mood-incongruence, increased with increasing PRS (OR=1.17, (bootstrapped p-value < .001, 95% C.I. 1.08 - 1.27)). Analyses excluding the SABD sample found comparable results (OR=1.20, bootstrapped p-value < .001, 95% C.I. 1.09, 1.32).

Discussion

Main Findings

Higher PRS-SCZ in BD [20,61] is well established. Here, we replicate and extend this observation, demonstrating a gradient of PRS associations across SCZ and BD subtypes (CLOZUK > SABD > BD I with psychosis > BD I without psychosis > BD II). We also show BD cases with psychosis carry a higher burden of SCZ risk-alleles, compared to BD without a history of psychosis. Furthermore, individuals with psychotic BD characterised by prominent mood-incongruent psychotic features, carry the highest burden of schizophrenia risk-alleles. There is a clear exposure-response gradient, with increasing PRS associated with psychotic BD and increasing mood-incongruence

289 (mood-incongruent > mood-congruent > no psychosis), supporting our hypothesis that
290 mood-incongruence indexes phenotypic features linked to SCZ liability.

291 Previously published work examining PRS for SCZ across BD, stratified by psychosis, did
292 not find significant discrimination [41,62] although a trend was observed, consistent with
293 the findings presented here. The most likely explanations for the enhanced signal in the
294 current analysis are: PRS were constructed using alleles derived from a larger SCZ-
295 GWAS discovery set which reduces measurement error plus improved power from both
296 this and the larger BD sample [63]. This group has shown[41], PRS-SCZ significantly
297 differentiate SABD from non-SABD subtypes, while finding no statistically significant
298 differential between BD stratified by psychosis, suggesting it is the nature of the
299 psychotic symptoms rather than their presence which better indexes liability shared
300 with SCZ. The current analysis supports this proposition that it is the level of mood-
301 incongruence rather than the presence of psychosis *per se* which better specifies a
302 shared biologically-validated dimensional trait, captured, but with less precision by the
303 SABD diagnostic category.

304 Psychosis and mood-incongruent psychotic features are known to be correlated to
305 poorer prognosis and treatment response[30-32] It is possible the trans-diagnostic
306 exposure-response gradient for PRS with the occurrence and nature of psychotic
307 symptoms presented here, could be the result of a general psychopathological factor
308 cutting across psychiatric disorders which influences the severity of psychopathology
309 generally, as well as, or rather than a psychosis-specific domain and that PRS derived
310 from SCZ GWAS may be indexing a general liability for psychopathology severity (at
311 least in part)[64] rather than a (SCZ) disease specific liability.

Implications

Our study supports the hypothesis that within BD, positive and disorganized psychotic symptoms, and in particular mood-incongruent psychotic features, represent a dimensionally defined stratum with underpinning biological validity. These features are not only phenotypically similar to those observed in prototypal schizophrenia but also index a greater shared genetic aetiology suggesting they share more pathophysiology [65]. It is notable that in those diagnosed with BD I with no history of psychosis, the association with schizophrenia liability was weaker but still on average higher than in the control group, while in the BD II subsample there was no overlap with SCZ liability. We are not suggesting psychotic features are the best or only index of shared pathophysiology, but having established stronger genetic links between the risk for SCZ and BD characterised by the occurrence of psychosis and level of mood-incongruence, we now have a basis to refine this signal. These findings represent a step towards the goal of reconceptualising phenotypic definitions using richer clinical signatures, measured across quantitative/qualitative domains including, symptom loadings and biomarker expression, outlined in the rationale for the Research Domain Criteria (RDoC) [66,67] and the road map for mental health research (ROAMER) [68] projects. It is probable however a multidimensional stratification process will harness the observed clinical heterogeneity better and define more precise patient-strata/subgroups in closer alignment with the underlying pathophysiology [68-70]

Methodological considerations

The phenotypic ratings used in the current analyses are based on both SCAN interviews and case-note review by raters with excellent inter-rater reliability, which is expected to minimise rates of missing data and reduce the likelihood of phenotypic

misclassification[71]. Our psychosis phenotypes are broadly defined and likely to represent imperfect measurements of a continuously distributed phenotype[72], imposing categorical constraints as we have done may reduce power. We generated PRS using a single discovery set p-value threshold < 0.05 and dealt with multiple comparisons, across different phenotypic categories/strata using bootstrap re-sampling approaches within each of our 4 independent analyses, adjusting for family-wise type-1 error proliferation using Bonferroni's correction. We have mitigated against potential confounding due to population stratification and potential batch effects across cases and controls, by partialling out the first 10 PCs and genotyping platforms from the PRS. The PRS were generated using most probable genotypes which can potentially reduce power due to a small (non-differential) loss of information at some markers making our results conservative, but the conclusions are unlikely to change. Finally, we have only examined the effect of common variants, as rare variants are not captured by current GWAS.

Conclusions

We show for the first time a gradient of polygenic liability across schizophrenia and bipolar disorder, indexed by the occurrence and level of mood-incongruence of positive and disorganised psychotic symptoms. This highlights the usefulness of genetic data to dissect clinical heterogeneity within and across disorders, and suggests further research could potentially aid in defining patient stratifiers with improved biological precision/validity, moving us tentatively towards precision medicine in psychiatry.

Acknowledgements

The work at Cardiff University was funded by Medical Research Council (MRC) Centre (MR/L010305/1) and Program Grants (G0800509). The CLOZUK sample was genotyped with funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 279227 (CRESTAR Consortium; http://www.crestar-project.eu/). . For the CLOZUK sample we thank Leyden Delta (Nijmegen, Netherlands) for supporting the sample collection, anonymisation and data preparation (particularly Marinka Helthuis, John Jansen, Karel Jollie and Anouschka Colson) and Andy Walker from Magna Laboratories (UK). We acknowledge Lesley Bates, Catherine Bresner and Lucinda Hopkins, at Cardiff University, for laboratory sample management.
We would like to acknowledge funding for BDRN from the Wellcome Trust and Stanley Medical Research Institute, all members of BDRN, and especially the participants who have kindly given their time to participate in our research.

Generation Scotland (GS) received core support from the Chief Scientist Office of the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council [HR03006]. Genotyping of the GS:SFHS samples was carried out by the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by the Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic Award "STratifying Resilience and Depression Longitudinally" (STRADL) Reference 104036/Z/14/Z)

The Type 1 Diabetes Genetics Consortium (T1DGC; EGA dataset EGAS00000000038) is a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and JDRF. Venous blood collection for the 1958 Birth Cohort (NCDS) was funded by the UK's Medical Research Council (MRC) grant G0000934, peripheral blood lymphocyte preparation by Juvenile Diabetes Research Foundation (JDRF) and WT and the cell-line production, DNA extraction and processing by WT grant 06854/Z/02/Z. Genotyping was supported by WT (083270) and the European Union (EU; ENGAGE: HEALTH-F4-2007- 201413).

The funders have had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication

V Escott-Price had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
J. Allardyce, G. Leonenko, A F J. Pardiñas, M. Hamshere and V. Escott-Price, all from MRC Centre for Neuropsychiatric Genetics and Genomics, Institute of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, UK conducted and are responsible for the data analysis.

Conflict of interest Disclosures: M.C. O'Donovan received a consultancy fee from Roche in July 2015. No other disclosures are reported

406 Reference List

407 1. *ICD-10 Classification of Mental and Behavioural Disorder:Clinical Descriptions and Diagnostic*
408 *Guidelines.* Geneva: World Health Organisation 1992.
409 2. *Diagnostic and Statistical Manual of Mental Disorders* 5th ed. Arlington: American
410 Psychiatric Association; 2013.
411 3. Spitzer RL, Endicott J, Robins E. Research diagnostic criteria: rationale and reliability.
412 *Archives of general psychiatry.* 1978;35(6):773-782.
413 4. Allardyce J, McCreadie RG, Morrison G, van Os J. Do symptom dimensions or categorical
414 diagnoses best discriminate between known risk factors for psychosis? *Social psychiatry and*
415 *psychiatric epidemiology.* 2007;42(6):429-437.
416 5. Derks EM, Allardyce J, Boks MP, Vermunt JK, Hijman R, Ophoff RA. Kraepelin was right: a
417 latent class analysis of symptom dimensions in patients and controls. *Schizophrenia bulletin.*
418 2012;38(3):495-505.
419 6. Dikeos DG, Wickham H, McDONALD C, et al. Distribution of symptom dimensions across
420 Kraepelinian divisions. *The British Journal of Psychiatry.* 2006;189(4):346-353.
421 7. Kitamura T, Okazaki Y, Fujinawa A, Yoshino M, Kasahara Y. Symptoms of psychoses. A factor-
422 analytic study. *The British Journal of Psychiatry.* 1995;166(2):236-240.
423 8. Lindenmayer J-P, Brown E, Baker RW, et al. An excitement subscale of the Positive and
424 Negative Syndrome Scale. *Schizophrenia research.* 2004;68(2):331-337.
425 9. McGorry P, Bell R, Dudgeon P, Jackson H. The dimensional structure of first episode
426 psychosis: an exploratory factor analysis. *Psychological medicine.* 1998;28(04):935-947.
427 10. McIntosh A, Forrester A, Lawrie S, et al. A factor model of the functional psychoses and the
428 relationship of factors to clinical variables and brain morphology. *Psychological medicine.*
429 2001;31(01):159-171.
430 11. Murray V, McKee I, Miller P, et al. Dimensions and classes of psychosis in a population
431 cohort: a four-class, four-dimension model of schizophrenia and affective psychoses.
432 *Psychological medicine.* 2005;35(04):499-510.
433 12. Ratakonda S, Gorman JM, Yale SA, Amador XF. Characterization of psychotic conditions: Use
434 of the domains of psychopathology model. *Archives of General Psychiatry.* 1998;55(1):75-81.
435 13. Serretti A, Olgiati P. Dimensions of major psychoses: a confirmatory factor analysis of six
436 competing models. *Psychiatry research.* 2004;127(1):101-109.
437 14. Alastair G. Cardno, Frühling V. Rijsdijk, Pak C. Sham, Robin M. Murray, Peter McGuffin. A
438 Twin Study of Genetic Relationships Between Psychotic Symptoms. *American Journal of*
439 *Psychiatry.* 2002;159(4):539-545.
440 15. Charney AW, Ruderfer DM, Stahl EA, et al. Evidence for genetic heterogeneity between
441 clinical subtypes of bipolar disorder. *Transl Psychiatry.* 2017;7:e993.
442 16. Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, Perlis RH. Genetic relationship between
443 five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet.* 2013;45:984-994.
444 17. Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, Sullivan PF. Common genetic
445 determinants of schizophrenia and bipolar disorder in Swedish families: a population-based
446 study. *Lancet.* 2009;373:234-239.
447 18. Maier W, Lichtermann D, Franke P, Heun R, Falkai P, Rietschel M. The dichotomy of
448 schizophrenia and affective disorders in extended pedigrees. *Schizophrenia*
449 *Research.*57(2):259-266.
450 19. Mortensen P, Pedersen CB, Melbye MM, Mors OO, Ewald HH. Individual and familial risk
451 factors for bipolar affective disorders in denmark. *Archives of General Psychiatry.*
452 2003;60(12):1209-1215.
453 20. Purcell SM, Wray NR, Stone JL, et al. Common polygenic variation contributes to risk of
454 schizophrenia and bipolar disorder. *Nature.* 2009;460(7256):748-752.

21. Consortium SWGotPG. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 2014;511:421-427.
22. Ruderfer D, Fanous A, Ripke S, McQuillin A, Amdur R. Cross-Disorder Working Group of Psychiatric Genomics Consortium, et al. Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. *Mol Psychiatry.* 2014;19(9):1017-1024.
23. Ford JM, Morris SE, Hoffman RE, et al. Studying hallucinations within the NIMH RDoC framework. *Schizophrenia bulletin.* 2014:sbu011.
24. Manchia M, Cullis J, Turecki G, Rouleau GA, Uher R, Alda M. The impact of phenotypic and genetic heterogeneity on results of genome wide association studies of complex diseases. *PloS one.* 2013;8(10):e76295.
25. Wray NR, Maier R. Genetic basis of complex genetic disease: the contribution of disease heterogeneity to missing heritability. *Current Epidemiology Reports.* 2014;1(4):220-227.
26. Diefendorf AR, Kraepelin E. Clinical psychiatry: A textbook for students and physicians, abstracted and adapted from the 7th German edition of Kraepelin's Lehrbuch der Psychiatrie In:1907:562.
27. Organization WH. Report of the international pilot study of schizophrenia. 1973.
28. Azorin JM, Akiskal H, Hantouche E. The mood-instability hypothesis in the origin of mood-congruent versus mood-incongruent psychotic distinction in mania: validation in a French National Study of 1090 patients. *Journal of affective disorders.* 2006;96(3):215-223.
29. Black DW, Nasrallah A. Hallucinations and delusions in 1,715 patients with unipolar and bipolar affective disorders. *Psychopathology.* 1989;22(1):28-34.
30. Goes FS, Zandi PP, Miao K, et al. Mood-incongruent psychotic features in bipolar disorder: familial aggregation and suggestive linkage to 2p11-q14 and 13q21-33. *American Journal of Psychiatry.* 2007;164(2):236-247.
31. Tohen M, Tsuang MT, Goodwin DC. Prediction of outcome in mania by mood-congruent or mood-incongruent psychotic features. *American Journal of Psychiatry.* 1992;149:1580-1580.
32. Toni C, Perugi G, Mata B, Madaro D, Maremmani I, Akiskal HS. Is mood-incongruent manic psychosis a distinct subtype? *European archives of psychiatry and clinical neuroscience.* 2001;251(1):12-17.
33. Green EK, Raybould R, Macgregor S, et al. Operation of the schizophrenia susceptibility gene, neuregulin 1, across traditional diagnostic boundaries to increase risk for bipolar disorder. *Archives of General Psychiatry.* 2005;62(6):642-648.
34. Park N, Juo SH, Cheng R, et al. Linkage analysis of psychosis in bipolar pedigrees suggests novel putative loci for bipolar disorder and shared susceptibility with schizophrenia. *Mol Psychiatry.* 2004;9(12):1091-1099.
35. Potash JB, Zandi PP, Willour VL, et al. Suggestive linkage to chromosomal regions 13q31 and 22q12 in families with psychotic bipolar disorder. *American Journal of Psychiatry.* 2003;160(4):680-686.
36. Raybould R, Green EK, MacGregor S, et al. Bipolar disorder and polymorphisms in the dysbindin gene (DTNBP1). *Biological psychiatry.* 2005;57(7):696-701.
37. Dudbridge F. Polygenic Epidemiology. *Genetic Epidemiology.* 2016;40(4):268-272.
38. Pasaniuc B, Price AL. Dissecting the genetics of complex traits using summary association statistics. *Nature Reviews Genetics.* 2016.
39. Wray NR, Lee SH, Mehta D, Vinkhuyzen AA, Dudbridge F, Middeldorp CM. Research review: Polygenic methods and their application to psychiatric traits. *Journal of child psychology and psychiatry, and allied disciplines.* 2014;55(10):1068-1087.
40. Consortium C-DGotPG. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature genetics.* 2013;45(9):984-994.
41. Hamshere ML, O'Donovan MC, Jones IR, et al. Polygenic dissection of the bipolar phenotype. *The British Journal of Psychiatry.* 2011;198(4):284-288.

42. Green EK, Hamshere M, Forty L, Gordon-Smith K, Fraser C, Russell E. Replication of bipolar disorder susceptibility alleles and identification of two novel genome-wide significant associations in a new bipolar disorder case-control sample. *Mol Psychiatry.* 2013;18:1302-1307.
43. Green EK, Rees E, Walters JTR, et al. Copy number variation in bipolar disorder. *Mol Psychiatry.* 2016;21(1):89-93.
44. Wing JK, Babor TT, Brugha TT, et al. Scan: Schedules fonr clinical assessment in neuropsychiatry. *Archives of General Psychiatry.* 1990;47(6):589-593.
45. Jones L, Scott J, Haque S, et al. Cognitive style in bipolar disorder. *The British Journal of Psychiatry.* 2005;187(5):431-437.
46. McGuffin P, Farmer A, Harvey I. A polydiagnostic application of operational criteria in studies of psychotic illness: Development and reliability of the opcrit system. *Archives of General Psychiatry.* 1991;48(8):764-770.
47. Leckman JF, Sholomskas D, Thompson D, Belanger A, Weissman MM. Best estimate of lifetime psychiatric diagnosis: a methodological study. *Archives of General Psychiatry.* 1982;39(8):879-883.
48. Pardiñas AF, Holmans P, Pocklington AJ, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and maintained by background selection. *bioRxiv.* 2016.
49. Hilner JE, Perdue LH, Sides EG, et al. Designing and implementing sample and data collection for an international genetics study: the Type 1 Diabetes Genetics Consortium (T1DGC). *Clinical trials.* 2010;7(1_suppl):S5-S32.
50. Amador C, Huffman J, Trochet H, et al. Recent genomic heritage in Scotland. *BMC genomics.* 2015;16(1):437.
51. Zuvich RL, Armstrong LL, Bielinski SJ, et al. Pitfalls of merging GWAS data: lessons learned in the eMERGE network and quality control procedures to maintain high data quality. *Genetic epidemiology.* 2011;35(8):887-898.
52. Delaneau O, Zagury J-F, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nature methods.* 2013;10(1):5-6.
53. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature genetics.* 2012;44(8):955-959.
54. McClellan J, McCurry C. Early onset psychotic disorders: diagnostic stability and clinical characteristics. *European child & adolescent psychiatry.* 1999;8(1):S13-S19.
55. Vieta E, Gasto C, Otero A, Nieto E, Vallejo J. Differential features between bipolar I and bipolar II disorder. *Comprehensive psychiatry.* 1997;38(2):98-101.
56. Craddock N, Jones I, Kirov G, Jones L. The Bipolar Affective Disorder Dimension Scale (BADDS)–a dimensional scale for rating lifetime psychopathology in bipolar spectrum disorders. *BMC psychiatry.* 2004;4(1):19.
57. Scott-Long JFJ. Models for nominal outcomes with case-specific data. In: *Regression Models for Categorical Dependent Variables Using STATA.* Texas 77845: Stata Press; 2006:252-254.
58. Westfall PHY, S.S,. *Resampling-based multiple testing: examples and methods for p-value adjustment.* New Jersey: John Wiley & Sons; 1993.
59. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet.* 2011;43:977-983.
60. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience.* 2015;4(1):7.
61. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet.* 2013;381:1371-1379.

554 62. Ruderfer DM, Fanous AH, Ripke S, McQuillin A, Amdur RL. Polygenic dissection of diagnosis
555 and clinical dimensions of bipolar disorder and schizophrenia. *Mol Psychiatry.* 2014;19:1017-
556 1024.
557 63. Dudbridge F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet.*
558 2013;9(3):e1003348.
559 64. Caspi A, Houts RM, Belsky DW, et al. The p Factor: One General Psychopathology Factor in
560 the Structure of Psychiatric Disorders? *Clinical psychological science : a journal of the*
561 *Association for Psychological Science.* 2014;2(2):119-137.
562 65. Keshavan MS, Morris DW, Sweeney JA, et al. A dimensional approach to the psychosis
563 spectrum between bipolar disorder and schizophrenia: the Schizo-Bipolar Scale.
564 *Schizophrenia research.* 2011;133(1):250-254.
565 66. Insel T, Cuthbert B, Garvey M, et al. Research domain criteria (RDoC): toward a new
566 classification framework for research on mental disorders. In: Am Psychiatric Assoc; 2010.
567 67. Insel TR. The NIMH research domain criteria (RDoC) project: precision medicine for
568 psychiatry. *American Journal of Psychiatry.* 2014;171(4):395-397.
569 68. Schumann G, Binder EB, Holte A, et al. Stratified medicine for mental disorders. *European*
570 *Neuropsychopharmacology.* 2014;24(1):5-50.
571 69. Insel TR, Cuthbert BN. Brain disorders? Precisely. *Science.* 2015;348(6234):499-500.
572 70. Joyce DW, Kehagia AA, Tracy DK, Proctor J, Shergill SS. Realising stratified psychiatry using
573 multidimensional signatures and trajectories. *Journal of Translational Medicine.*
574 2017;15(1):15.
575 71. Allardyce J, Morrison G, Van Os J, Kelly J, Murray RM, McCreadie RG. Schizophrenia is not
576 disappearing in south-west Scotland. *The British journal of psychiatry : the journal of mental*
577 *science.* 2000;177:38-41.
578 72. Van Os J, Linscott RJ, Myin-Germeys I, Delespaul P, Krabbendam L. A systematic review and
579 meta-analysis of the psychosis continuum: evidence for a psychosis proneness–persistence–
580 impairment model of psychotic disorder. *Psychological medicine.* 2009;39(02):179-195.

581

Table 1:Differential Association of PRS across variously defined BD strata (controls as comparator category)

	N (subsample)	RR	Bootstrapped p-value	Bonferroni Corrected p-value	Bootstrapped 95% confidence intervals
CLOZUK	4,976	1.94	< .001	< .001	1.86, 2.01
		A) Bipolar Disorder cases stratified by RDC defined subtypes			
SABD	356	1.37	< .001	< .001	1.22, 1.54
BD I	2,775	1.30	< .001	< .001	1.24. 1.36
BD II	1,268	1.04	0.26	0.26	0.97, 1.11
		B) Bipolar Disorder cases stratified by lifetime occurrence of psychosis			
No LEP	2,079	1.09	0.001	0.004	1.04, 1.15
LEP	2,296	1.36	< .001	< .001	1.29, 1.43
		C) Psychotic Bipolar Disorder cases stratified by level of mood incongruence			
Low LMI	1,126	1.24	< .001	< .001	1.17, 1.33
High LMI	981	1.46	< .001	< .001	1.36, 1.57
		D) Sensitivity Analysis: Psychotic Bipolar Disorder cases stratified by levels of mood incongruence (excluding SABD cases)			
Low LMI	1,068	1.25	< .001	< .001	1.16, 1.33
High LMI	699	1.49	< .001	< .001	1.37, 1.62

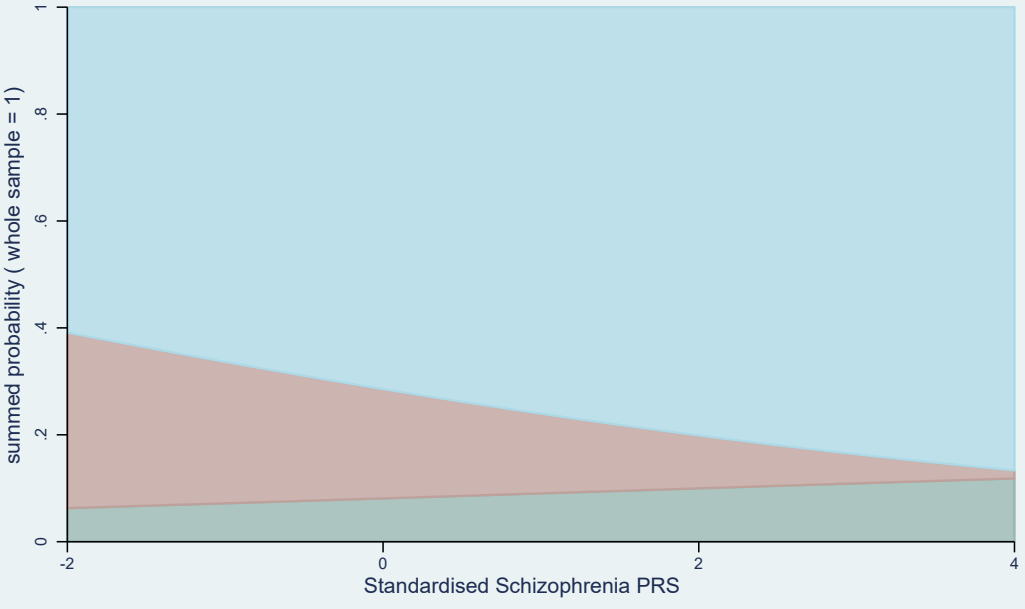
CLOZUK – Treatment resistant Schizophrenia, treated with clozapine, BD I - RDC bipolar disorder subtype I, BD II - RDC bipolar disorder type II, SABD-RDC bipolar schizoaffective disorder, LEP – lifetime ever occurrence of psychotic symptoms, LMI – lifetime pattern of low/high mood incongruent psychotic features RR – relative risk ratio PRS adjusted for 1st 10 PCs and genotyping platform

Table 2: PRS-SCZ associations among cases

	RR	Bootstrapped p-value	Bonferroni corrected p-value	Bootstrapped 95% C.I.
SABD compared to TRS	0.71	< .001	< .001	0.63, 0.80
BD I compared to TRS	0.67	< .001	< .001	0.64, 0.71
BD II compared to TRS	0.54	< .001	< .001	0.50, 0.57
SABD compared to BD II	1.32	< .001	< .001	1.16, 1.50
BP I compared to BD II	1.25	< .001	< .001	1.16, 1.35
SABD compared to BD I	1.05	0.41	0.41	0.93, 1.18

TRS - treatment resistant schizophrenia, treated with clozapine, BDI - RDC bipolar disorder subtype I, BD II - RDC bipolar disorder type II, SABD-RDC bipolar schizoaffective disorder, RR – relative risk ratio PRS adjusted for 1st 10 PCs and genotyping platform 95% bootstrapped C.I. - 95% confidence intervals.

Figure 1. Probability of RDC bipolar subtype as a function of polygenic risk scores (PRS) for schizophrenia



x-axis - standardised PRS in standard deviation units, SABD - schizoaffective bipolar type, BD I - bipolar disorder type I , BD II - bipolar type II

Figure 2: Percentage of bipolar subtype as a function of PRS for schizophrenia - grouped by decile

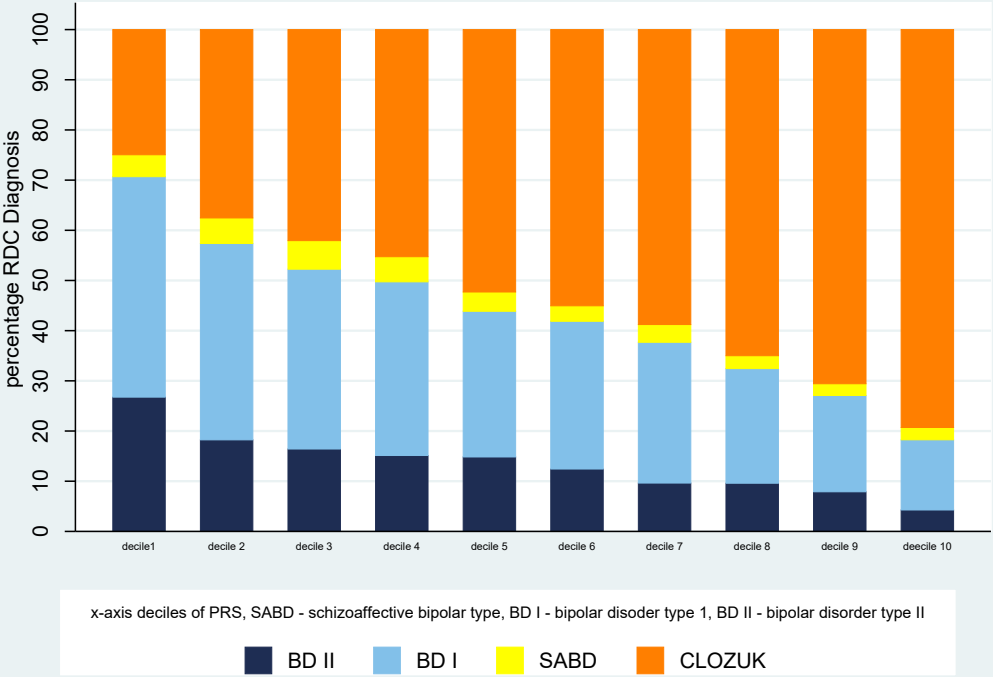


Figure 2. Relative Risk Ratio of PRS with subtypes of BD compared with controls (CLOZUK included for comparison)