Genetic association study of psychotic experiences in UK Biobank

Sophie E. Legge[1], Hannah J. Jones[2,3,4], Kimberley M. Kendall[1], Antonio F. Pardiñas[1], Georgina Menzies[5], Mathew Bracher-Smith[1], Valentina Escott-Price[1,5], Elliott Rees[1], Katrina A.S. Davis[6], Matthew Hotopf[6], Jeanne E. Savage[8], Danielle Posthuma[8], Peter Holmans[1], George Kirov[1], Michael J. Owen[1], Michael C. O'Donovan[1], Stanley Zammit[1,2,4*], James T.R. Walters[1*]

* Both authors contributed equally to this project

Corresponding author

Professor James Walters

Email: WaltersJT@Cardiff.ac.uk

MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, UK

Affiliations

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

[2] Centre for Academic Mental Health, Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

[3] Medical Research Centre, Integrative Epidemiology Unit, University of Bristol, Bristol, UK

[4] National Institute for Health Research Bristol Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust, University of Bristol, Bristol, UK

[5] UK-Dementia Research Institute at Cardiff, Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, UK

[6] Institute of Psychiatry Psychology and Neuroscience, King's College London, South London and Maudsley NHS Foundation Trust and NIHR Biomedical Research Centre, London, UK

[7] South London and Maudsley NHS Foundation Trust, UK

[8] Department of Complex Trait Genetics, Centre for Neurogenomics and Cognitive Research, Amsterdam Neuroscience, VU University, Amsterdam, The Netherlands

Abstract

Psychotic experiences, such as hallucinations and delusions, are reported by approximately 5%-10% of the general population, though only a small proportion of individuals develop psychotic disorders such as schizophrenia or bipolar disorder. Studying the genetic aetiology of psychotic experiences in the general population, and its relationship with the genetic aetiology of other disorders, may increase our understanding of their pathological significance. Using the population-based UK Biobank sample, we performed the largest genetic association study of psychotic experiences in individuals without a psychotic disorder. We conducted three genome-wide association studies (GWAS) for (i) any psychotic experience (6123 cases vs. 121,843 controls), (ii) distressing psychotic experiences (2143 cases vs. 121,843 controls), and (iii) multiple occurrence psychotic experiences (3337 cases vs. 121,843 controls). Analyses of polygenic risk scores (PRS), genetic correlation, and copy number variation (CNV) were conducted to assess whether genetic liability to psychotic experiences is shared with schizophrenia and/or other neuropsychiatric disorders and traits. GWAS analyses identified four loci associated with psychotic experiences including a locus in Ankyrin-3 (*ANK3*, OR=1.16, *p*=3.06 x 10^{-8}) with any psychotic experience and a locus in cannabinoid receptor 2 gene (*CNR2,* OR=0.66, *p*=3.78x10^{-8}) with distressing psychotic experiences. PRS analyses identified associations between psychotic experiences and genetic liability for schizophrenia, major depressive disorder, and bipolar disorder, and these associations were stronger for distressing psychotic experiences. Genetic correlation analysis identified significant genetic correlations between psychotic experiences and major depressive disorder, schizophrenia, autism spectrum disorder and a cross-disorder GWAS. Individuals reporting psychotic experiences had an increased burden of CNVs previously associated with schizophrenia (OR=2.04, *p*=2.49x10^{-4}) and of those associated with neurodevelopmental disorders more widely (OR=1.75, *p*=1.41x10^{-3}). In conclusion, we identified four genome-wide significant loci in the largest GWAS of psychotic experiences from the population-based UK Biobank sample and found support for a shared genetic aetiology between psychotic experiences and schizophrenia, but also major depressive disorder, bipolar disorder and neurodevelopmental disorders.

Introduction

Psychotic experiences, such as hallucinations and delusions, are features of psychiatric disorders, for example schizophrenia and bipolar disorder, but they are also reported by approximately 5%-10% of the general population[1,2]. Psychotic experiences are only considered to be symptoms of psychiatric illness if they co-occur with other features of that disorder, including some aspect of psychosocial impairment. It is currently unclear whether psychotic experiences in the general population are (i) on a spectrum that at the extreme relates specifically to schizophrenia, (ii) largely unrelated to the psychotic symptoms experienced in schizophrenia and other major mental disorders, or (iii) related to liability to major mental disorders more generally.

Twin studies and genome-wide association studies (GWAS) have provided evidence that psychotic experiences are heritable (30-50% from twin studies[3-5], 3-17% for SNP-heritability estimates[6,7]), indicating that common genetic variants play a role in their aetiology. There have been three GWAS of psychotic experiences to date, all conducted in adolescent samples[7-9] with relatively small sample sizes (largest total sample n=10,098) and no reported genome-wide significant findings. These studies have focused on adolescence under the hypothesis that psychotic experiences at this age could be associated with an increased risk of mental health disorders in later life[10]. Whilst there was an initial assumption that psychotic experiences in adolescence would specifically increase the risk for schizophrenia, evidence suggests a non-specific increased risk for a broader psychopathology[10], suggesting that research on psychotic experiences may have an important role in understanding the pathway to a wide array of clinical diagnoses[5]. However, to date no study has found strong evidence for association between genetic liabilities for schizophrenia or any other mental disorder with psychotic experiences[7,8,11-13].

Although many individuals with a lifetime history of psychotic experiences have their first experience in adolescence, nearly a quarter of first onset psychotic experiences occur after 40 years of age[14]. It is possible that psychotic experiences first occurring in adulthood have a different aetiology to those occurring in adolescence. Our aims were to use the UK Biobank to (i) identify genetic loci associated with psychotic experiences reported by adults in a population-based study, and (ii) to determine whether genetic liability to psychotic experiences is shared with schizophrenia and/or other neuropsychiatric disorders and traits.

3

Methods

Sample

Study individuals were from the UK Biobank, a large prospective population-based cohort study of approximately 500,000 individuals aged between 40-69 who were recruited from 22 assessment centres across the UK between 2006-2010[15]. The North West Multi-Centre Ethics Committee granted ethical approval to UK Biobank and all participants provided informed consent. This study was conducted under UK Biobank project numbers 13310 and 14421.

Psychotic experiences phenotypes

A Mental Health Questionnaire (MHQ) was sent to all participants in 2016-2017 who provided an email address and was completed by a total of 157,397 individuals (46% of those emailed, 31% of the total UK Biobank sample). For psychotic experiences, participants were asked about previous experience of visual hallucinations, auditory hallucinations, delusions of reference, delusions of persecution, how often these experiences occurred and how distressing they found them. The questions asked are detailed in **Supplementary Methods**. All analyses were conducted excluding individuals with a diagnosis of a psychotic disorder (schizophrenia, bipolar disorder or any other psychotic disorder) as defined in **Supplementary Methods.**

We selected three primary phenotypes for GWAS: (i) *any psychotic experience* defined as a positive response to any of the four symptom questions (UKB field IDs: 20471, 20463, 20474, 20468), (ii) a *distressing psychotic experience,* defined as *any psychotic experience* that was rated as 'a bit', 'quite' or 'very' distressing (UKB field ID: 20462), (iii) *multiple occurrence psychotic experiences* defined as *any psychotic experience* that occurred on more than one occasion (UKB field IDs: 20473, 20465, 20476, 20470). As a comparator group, we included individuals who provided a negative response to all four psychotic experience symptom questions. **Supplementary Figure 1** details the overlap in samples between these phenotypes.

We investigated additional psychotic experience phenotypes to those described above for association with PRS: (i) *visual hallucination* (UKB field ID: 20471), (ii) *auditory hallucination* (UKB field ID: 20463), (iii) *delusion of reference* (UKB field ID: 20474), (iv) *delusion of persecution* (UKB field ID: 20468), and (v) *non-distressing psychotic experiences* defined as any psychotic experience rated as not distressing (UKB field ID: 20462).

Genetic data

Genetic data for the study participants was provided by UK Biobank and the imputation and quality control procedures are fully described elsewhere[16]. The data release contained 488,377 participants assayed on either the UK Biobank Axiom or the UK BiLEVE Axiom purpose-built arrays at the Affymetrix Research Services Laboratory. Standard quality control procedures were applied prior to imputation using Haplotype Reference Consortium (HRC)[17] and UK10K haplotype reference[18] panels. We applied additional quality control filters to select high quality SNPs[19]; minor allele frequency (MAF) > 0.01, imputation score (INFO) > 0.8, missingness < 0.05, Hardy-Weinberg equilibrium (HWE) < 1×10^{-6}, and removed SNPs imputed by the UK10K haplotype reference[18] dataset in accordance with guidance from the UK Biobank (URL: http://www.ukbiobank.ac.uk/2017/07/important-note-about-imputed-genetics-data/). One member from each related pair with a kinship coefficient > 0.15 was excluded from analyses, preferentially retaining individuals that had experienced a psychotic experience, and were otherwise removed at random.

Analyses were restricted to individuals with a self-reported British and Irish background (UKB field ID: 21000) and principal components supplied by UK Biobank[16] (UKB field ID: 22009) were used to assess and control for population structure. European genetic ancestry was assessed using the 'covMCD' function in the R package 'robustbase'[20,21] and additional exclusions applied (described in **Supplementary Methods**).

GWAS analysis

To identify genetic risk variants for psychotic experiences, association analysis was performed in SNPTEST (v2.5.4)[22] using bgen v1.2 imputed dosage data[23]. Over 7.5 million SNPs were included in each GWAS. An additive logistic regression model was used including as covariates genotyping array, the top five principal components (as recommended for most GWAS approaches[24]), and any additional principal components from the first 20 that were nominally associated ($p < 0.05$) with the GWAS phenotype in a logistic regression. To obtain relatively independent index SNPs, linkage disequilibrium (LD) clumping was performed in PLINK[25] ($r^2 < 0.1$, $p < 1 \times 10^{-4}$, window size < 3MB) for each GWAS using a reference panel of 1000 randomly selected UK Biobank individuals with confirmed European ancestry. Functional annotation was conducted using FUMA[26]. LDSC[27] was used to calculate the LD score intercept and heritability on the observed scale using the summary statistics from each psychotic experience GWAS.

Validation analyses in ALSPAC cohort

To assess the reproducibility of the psychotic experience GWAS, we used the summary statistics from the *any psychotic experience* GWAS to target psychotic experiences in the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort[28,29], which have been previously described[11,30]. A psychotic experience PRS was generated for each ALSPAC participant using the accepted method [31] and logistic regression was used to test for association between PRS and psychotic experiences reported at age 12 and 18 years. See **Supplementary Methods** for further details.

Polygenic risk scores

To examine the relationship between psychotic experiences and genetic risk for psychiatric and personality traits from publicly available GWAS datasets (that did not include UK Biobank where possible) included; schizophrenia[32], bipolar disorder[33], major depressive disorder[34], neuroticism[35], and intelligence[36]. PRSs were generated using the method described by the PGC[31] and detailed in **Supplementary Methods**. The intelligence GWAS excluded UK Biobank participants (n=74,214 individuals remaining, summary statistics specifically derived for this study). As it was not possible to obtain the major depressive disorder GWAS summary statistics excluding UK Biobank participants, we conducted sensitivity analyses removing individuals with a diagnosis of depression to control for overlap with the discovery dataset. The primary analysis used standardised scores generated from SNPs with a discovery sample p-value threshold of $p \leq 0.05$, but associations at 10 other p-value thresholds were also tested. A logistic regression model was used to test the association of each PRS with various psychotic experience phenotypes, covarying for the first five principal components and genotyping array.

Genetic correlations

LDSC[27,37] was used to calculate the genetic correlation between each psychotic experience GWAS and psychiatric and personality traits. External GWAS datasets used to generate the correlations were the same as those used for PRS analysis (detailed above) but also included attention deficit/hyperactivity disorder (ADHD)[38], autism spectrum disorder[39], and the PGC cross-disorder analysis[40]. A Bonferroni correction was applied to control for multiple testing.

Copy number variation

CNV calling, which has been described in detail elsewhere[41], was carried out using biallelic markers common to both genotyping platforms using PennCNV-Affy[42]. Exclusion criteria included ≥ 30 CNVs, waviness factor <-0.03 or > 0.03, call rate < 0.96 for individual samples, LRR SD > 0.35 and coverage of < 20 probes, density coverage of < 1 probe per 20,000 base pairs or a confidence score of < 10 for individual CNVs. We compared carrier status of rare CNVs previously associated with (i) schizophrenia[43] and (ii) neurodevelopmental disorders more widely[44] (which includes all schizophrenia-associated CNVs and excludes 15q11.2 duplication because of its high frequency following our previous publications[41]) with the three primary psychotic experience phenotypes used for GWAS. Association analyses were carried out using logistic regression and included age, sex and genotyping array as covariates.

Results

A total of 7803 (5.0%, 60.0% female) individuals from UK Biobank who completed the MHQ reported at least one psychotic experience, 3012 individuals rated the psychotic experience as distressing, and a total of 4388 reported multiple occurrences of at least one psychotic experience. A total of 147,461 individuals (56.0% female) who reported no psychotic experiences constituted the comparison group for our association analyses. The mean age of the first psychotic experience was 31.6 years (SD=17.6) (**Supplementary Figure 4**) with 2341 (35.2%) first occurring before the age of 20 or for as long as the participant could remember, 2137 (32.1%) between the ages of 20-39, and 2176 (32.7%) between the ages of 40 and 76. We excluded 198 individuals who had a diagnosis of schizophrenia, 818 with bipolar disorder, and 346 with other psychotic disorders from all genetic analyses (detailed in **Supplementary Methods)**.

GWAS

The GWAS of *any psychotic experience* in 6123 cases and 121,843 controls (following QC, exclusions detailed in **Supplementary Methods**) identified two variants that were associated at the genome-wide significance (GWS) level of $p < 5x10^{-8}$ (**Figure 1**, **Table 1**, λ_{GC}=1.05, LD Score regression intercept=1.00); rs10994278, an intronic variant within Ankyrin-3 (*ANK3*) on chromosome 10 (OR=1.16, 95% confidence intervals (CI)=1.10-1.23, p=3.06x10^{-8}), and rs549656827, an intergenic variant on chromosome 5 (OR=0.61, 95% CI=0.50-0.73, p=3.30x10^{-8}).

7

A second GWAS restricting the cases to 2143 individuals with *distressing psychotic experiences* identified two GWS variants (**Table 1**, λ_{GC}=1.03, LD Score intercept=1.01); rs75459873 intronic to cannabinoid receptor 2 (*CNR2*) on chromosome 1 (OR=0.66, 95% CI=0.56-0.78, p=3.78x10^{-8}) and rs3849810, an intergenic variant on chromosome 8 (OR=1.22, 95% CI=1.13-1.31, p=4.55x10^{-8}).

The third GWAS restricting the cases to 3337 individuals who reported *multiple occurrences of psychotic experiences* did not identify any associated variants at GWS (λ_{GC}=1.03, LD Score intercept=1.00).

Summary statistics for SNPs associated at p<1x10^{-5} for each GWAS are given in **Supplementary Tables 1-3** and QQ plots are displayed in **Supplementary Figure 5**. SNP based heritability estimates calculated by LDSC[27] for the GWAS of *any psychotic experience* was h^2=1.71% (95% CI=1.02-2.40%). The other GWAS analyses did not requirements (case n>5,000 *and* Z-score>4[45]) for heritability or genetic correlation analyses with LDSC.

Validation analyses in ALSPAC

There was evidence of association between the PRS calculated using the *any psychotic experience* GWAS from UK Biobank at the p-value threshold of ≤0.5 and definite psychotic experiences between ages 12 and 18 years in ALSPAC (OR=1.13; 95% CI=1.02-1.25; R^2=0.002; p=0.02). This finding was consistent for thresholds below p<0.05 and when using measures from age 18 years only (**Supplementary Figure 6**, all results detailed in **Supplementary Table 4**). However, the psychotic experiences PRS was also associated with the presence of major depressive disorder at age 18 (OR = 1.19; 95% CI=1.05-1.35; R^2=0.004; p=0.01).

Relationship between association at *CNR2* and cannabis use

We investigated the possibility of either a mediating or moderating effect of cannabis use on the association between *distressing psychotic experiences* and rs75459873 at the cannabinoid receptor gene *CNR2*. Cannabis use itself (UKB field id: 20453) was significantly associated with *distressing psychotic experiences* (OR=1.36, 95% CI=1.32-1.40, p=9.16 x 10^{-88}), but rs75459873 was not associated with cannabis use (OR=0.99, 95% CI=0.95-1.04, p=0.70). Furthermore, the association between rs75459873 and *distressing psychotic experiences* (OR=0.62, 95% CI=0.52-0.75) was unchanged in a model including cannabis use as a covariate (OR=0.62, 95% CI=0.52-0.75) or as an interaction term (OR=0.59, 95% CI=0.48-0.73, interaction p=0.31). Thus, we found no indication of either a mediating or moderating effect of cannabis use on the association between rs75459873 and *distressing psychotic experiences*.

Polygenic risk score analysis

We found evidence of a weak association between *any psychotic experience* and genetic liability indicated by PRS at a $p<0.05$ threshold for SNP inclusion for schizophrenia (OR=1.09; 95% CI=1.06-1.12; adjusted R^2=0.001; p=2.96x10^{-11}), major depressive disorder (OR=1.16; 95% CI=1.13-1.19; R^2=0.003; p=1.48x10^{-30}), and bipolar disorder (OR=1.07; 95% CI=1.04-1.10; R^2=0.001; p=5.11x10^{-7}). These findings were consistent across most p-value thresholds (**Figure 2** and **Supplementary Figure 7**). **Supplementary Figure 8** compares the average PRS scores (SNP inclusion threshold $p<0.05$) for each psychotic experience phenotype. Individuals with *distressing psychotic experiences* had significantly higher PRS scores for schizophrenia, bipolar disorder and major depressive disorder than those with *non-distressing psychotic experiences* (**Supplementary Table 6**).

We also considered individual psychotic symptoms and found that PRS for schizophrenia was more strongly associated with delusions of persecution than the other psychotic symptoms measured (**Supplementary Table 6**). This pattern was similar for PRS related to bipolar disorder and major depressive disorder and was consistent across p-value thresholds (**Supplementary Figure 9-10**). The association with psychotic experience phenotypes for bipolar disorder PRS and major depressive disorder PRS remained significant in analyses controlling for schizophrenia PRS (at $p<0.05$ threshold, **Supplementary Table 7**). The association with major depressive disorder PRS remained significant when individuals with a diagnosis of depression were removed to control for any potential overlap with the discovery dataset (**Supplementary Table 8**).

Genetic correlations

Significant genetic correlations (r_g) were observed between *any psychotic experience* and major depressive disorder (r_g=0.46, p=4.64x10^{-11}), autism spectrum disorder (r_g=0.39, p=1.68x10^{-4}), the PGC cross-disorder GWAS (r_g=0.30, p=1.53x10^{-3}) and schizophrenia (r_g=0.21, p=7.29x10^{-5}). **Figure 3** displays these genetic correlations and **Supplementary Table 5** details the full results. The other GWAS analyses did not meet requirements[45] for genetic correlation analysis.

Copy number variation

Individuals reporting *distressing psychotic experiences* in particular had an increased burden of CNVs previously associated with schizophrenia (OR=2.04; 95% CI=1.39-2.98; p=2.49x10^{-4}) and neurodevelopmental disorders (OR=1.75; 95% CI=1.24-2.48; p=1.41x10^{-3}). There was evidence

of an increased burden of these CNVs in individuals reporting *any psychotic experience* but not *multiple occurrence psychotic experiences* (**Table 2**).

Discussion

We conducted the largest GWAS of psychotic experiences using the population-based UK Biobank sample, and identified four genome-wide significant loci; an intronic variant to *ANK3*, an intronic variant to *CNR2,* and two intergenic loci. PRS, genetic correlation and CNV analyses indicated a shared genetic aetiology between psychotic experiences and schizophrenia, but also neurodevelopmental disorders and other psychiatric disorders such as major depressive disorder and bipolar disorder. The shared genetic aetiology between psychotic experiences and major mental health conditions was stronger for psychotic experiences rated as distressing.

The primary GWAS findings are related to intronic variants in *ANK3* and *CNR2*. The GWAS of *any psychotic experience* identified two significant loci, the most significant of which was indexed by rs10994278, an intronic variant to *ANK3* (OR=1.16, p=3.06x10^{-8}). The *ANK3* gene encodes ankyrin-G, a protein that has been shown to regulate the assembly of voltage-gated sodium channels and is essential for normal synaptic function[46]. *ANK3* is one of strongest and most replicated genes for bipolar disorder[33], and variants within *ANK3* have also been associated in the PGC cross-disorder GWAS[40], and in a rare variant analysis of autism spectrum disorder[47].

The GWAS of *distressing psychotic experiences* also identified two significant loci, the most significant of which was indexed by rs75459873, an intronic variant to *CNR2* (OR=0.66, p=3.78x10^{-8}). *CNR2* encodes for CB2, one of two well-characterised cannabinoid receptors (CB1 being the other). CB2 receptors they have abundant expression in the immune system, microglia and some neurons[48]. Several lines of evidence have implicated the endocannabinoid system in psychiatric disorders including schizophrenia[49,50] and depression[51]. The main psychoactive agent of cannabis, Δ^9-tetrahydrocannabinol, can cause acute psychotic symptoms and cognitive impairment[52]. Given that cannabis use is strongly associated with psychotic experiences, we tested, but found no evidence for, a mediating or moderating effect of cannabis use on the association of rs75459873 and *distressing psychotic experiences.* However, whilst no evidence was found in this study, a mediating effect of cannabis use cannot be ruled out given there may be questionable reliability in self-reported illicit cannabis use.

We found that psychotic experiences shared genetic risk with several psychiatric disorders, which included, but was not specific to, schizophrenia. PRS analyses identified associations between psychotic experiences and genetic liability for schizophrenia, major depressive disorder and bipolar disorder but not for intelligence or neuroticism. The magnitude of associations in the PRS analysis was comparable between psychiatric disorders and was very small (maximum r^2 value<1% and AUC<0.55). We found particular enrichment of schizophrenia, major depressive disorder and bipolar disorder PRS in cases that found the psychotic experiences distressing and for delusions of persecution. Genetic correlation analysis identified significant genetic correlations between psychotic experiences and major depressive disorder (r_g=0.46), schizophrenia (r_g=0.21), autism spectrum disorder (r_g=0.39), and for the PGC cross-disorder GWAS (r_g=0.30). Furthermore, we found an increased burden of CNVs previously associated with schizophrenia (OR=2.04) and neurodevelopmental disorders more widely (OR=1.75) in individuals with distressing psychotic symptoms. There was also an increased burden of CNVs in those reporting any psychotic experience but the association was stronger for distressing psychotic experiences. All schizophrenia-associated CNVs are also associated with neurodevelopmental disorders such as intellectual disability and autism spectrum disorder, and in fact penetrance is higher in these disorders[43]. Furthermore, CNVs in UK Biobank have been associated with a range of outcomes including cognitive performance[41,53] and depression[54], adding strength to our findings of a lack of specificity for psychotic experiences genetic risk.

A number of studies have demonstrated that psychopathology in the population is best described by a bifactor model with a common latent trait as well as specific traits, and that psychotic experiences index the more severe end of the common or shared trait[56,57]. Our findings of non-specificity of genetic risk for psychotic experiences with risk for other disorders are consistent with previous studies[7,8,11,12]. Nonetheless, despite lacking specificity, our results suggest that incorporating questions about frequency and, in particular, distress of psychotic experiences to self-reported assessments may allow a more valid identification of experiences that index schizophrenia and major mental health disorder liability.

Given the heritability estimate was very low (1.7%) for the *any psychotic experience* GWAS and the low amount of variance explained in our PRS analysis, and those of others[7,11], our findings indicate that understanding the genetics of psychotic experiences is unlikely to have an important impact on understanding the genetics of schizophrenia specifically. However, given that the small heritability of psychotic experiences is substantially shared with other psychiatric

disorders, and that non-genetic aetiology of psychotic experiences has been shown to be shared[55]; then studies of psychotic experience aetiology may be informative for psychiatric disorders in general.

Strengths and limitations

Strengths of this study include the large sample size (approximately 10x that of previous studies), which is required for genetic association studies, the use of an adult cohort, and the use of multiple psychotic experience phenotypes, which increase confidence in our findings. Furthermore, individuals with psychotic disorders were removed from our analyses, and thus the results will not be driven by this smaller subset of individuals.

One limitation of this study relates to the retrospective measurement of lifetime psychotic experiences by self-report from an online questionnaire, as this increases the likelihood of measurement error. A further limitation is the evidence of a 'healthy volunteer' bias for the participants recruited to UK Biobank and the sample cannot be therefore considered representative of the general population[58]. We also found that the participants that completed the MHQ had significantly higher intelligence and lower schizophrenia, depression and neuroticism PRS compared to UK Biobank participants who did not complete the MHQ (**Supplementary Table 9**). It is possible therefore that our results are affected by selection bias. Lastly, we were not able to entirely de-duplicate the UK Biobank individuals from all of the external datasets used for the polygenic risk score analysis. We excluded datasets that explicitly included the UK Biobank where possible, and failing that removed cases in sensitivity analyses.

Conclusions

In the largest GWAS of psychotic experiences from the population-based UK Biobank sample, we identified four genome-wide significant loci associated with psychotic experiences including an intronic variant to *ANK3*, an intronic variant to *CNR2,* and two intergenic loci. We found support for a shared genetic aetiology between psychotic experiences and several psychiatric disorders including schizophrenia, major depressive disorder, bipolar disorder, and neurodevelopmental disorders indicating that psychotic experiences are not specifically related to schizophrenia, but rather to a general risk for mental health disorders.
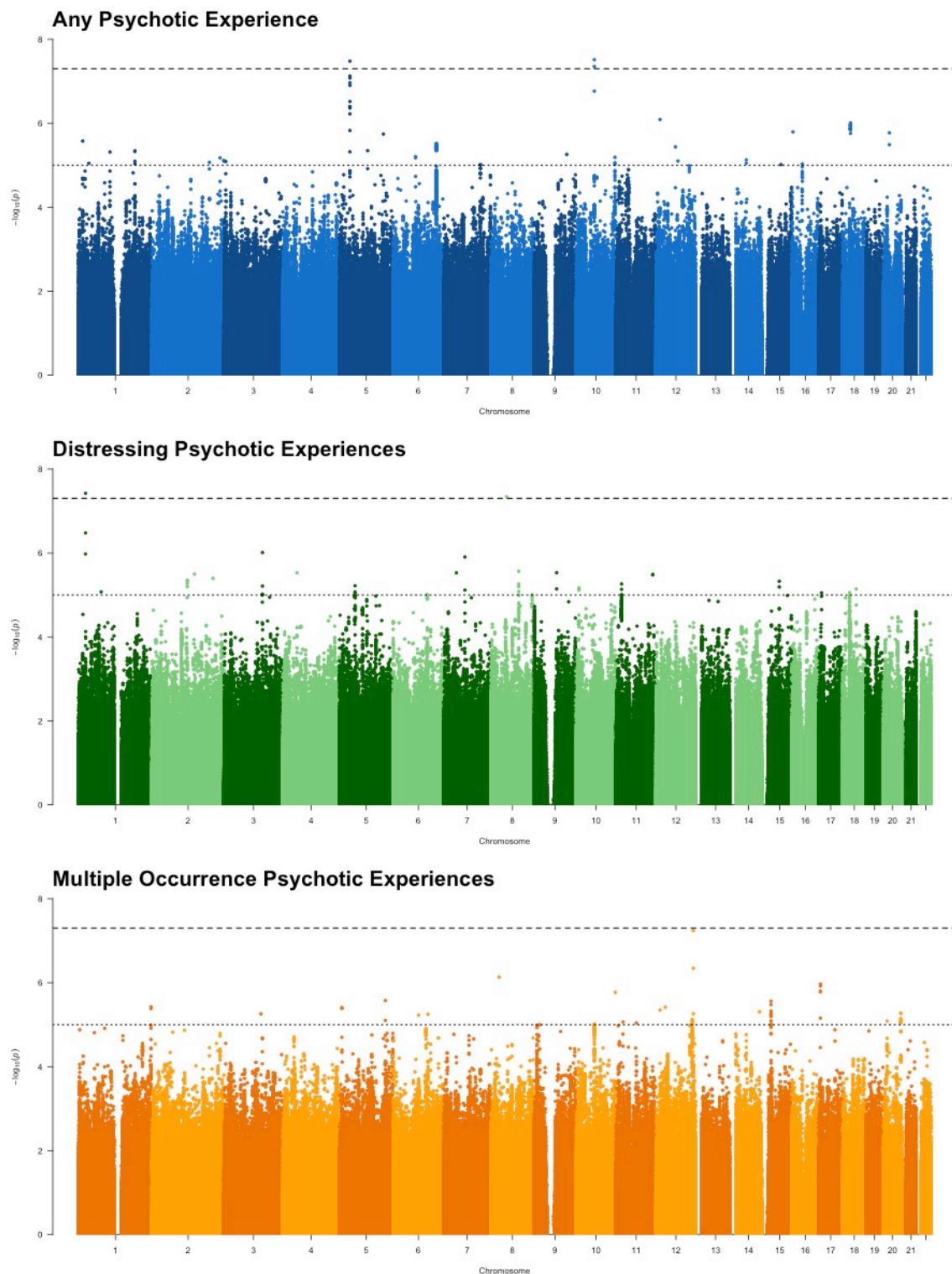
Acknowledgements

References

1. McGrath, J.J., Saha, S., Al-Hamzawi, A., Alonso, J., Bromet, E.J., Bruffaerts, R. *et al.* Psychotic Experiences in the General Population: A Cross-National Analysis Based on 31,261 Respondents From 18 Countries. 2015; **72**: 697-705.

2. van Os, J., Linscott, R.J., Myin-Germeys, I., Delespaul, P. & Krabbendam, L. A systematic review and meta-analysis of the psychosis continuum: evidence for a psychosis proneness-persistence-impairment model of psychotic disorder. *Psychol Med* 2009; **39**: 179-195.

3. Zavos, H.M., Freeman, D., Haworth, C.M., McGuire, P., Plomin, R., Cardno, A.G. *et al.* Consistent etiology of severe, frequent psychotic experiences and milder, less frequent manifestations: a twin study of specific psychotic experiences in adolescence. *JAMA Psych* 2014; **71**: 1049-1057.

4. Ronald, A. Recent quantitative genetic research on psychotic experiences: new approaches to old questions. *Curr Opin Behav Sci* 2015; **2**: 81-88.

5. Ronald, A. & Pain, O. A systematic review of genome-wide research on psychotic experiences and negative symptom traits: new revelations and implications for psychiatry. *Hum Mol Genet* 2018; **27**: R136-R152.

6. Sieradzka, D., Power, R.A., Freeman, D., Cardno, A.G., Dudbridge, F. & Ronald, A. Heritability of Individual Psychotic Experiences Captured by Common Genetic Variants in a Community Sample of Adolescents. *Behav Genet* 2015; **45**: 493-502.

7. Pain, O., Dudbridge, F., Cardno, A.G., Freeman, D., Lu, Y., Lundstrom, S. *et al.* Genome-wide analysis of adolescent psychotic-like experiences shows genetic overlap with psychiatric disorders. *Am J Med Genet B* 2018; **177**: 416-425.

8. Zammit, S., Hamshere, M., Dwyer, S., Georgiva, L., Timpson, N., Moskvina, V. *et al.* A population-based study of genetic variation and psychotic experiences in adolescents. *Schizophr Bull* 2014; **40**: 1254-1262.

9. Ortega-Alonso, A., Ekelund, J., Sarin, A.P., Miettunen, J., Veijola, J., Jarvelin, M.R. *et al.* Genome-Wide Association Study of Psychosis Proneness in the Finnish Population. *Schizophr Bull* 2017; **43**: 1304-1314.

10. Fisher, H.L., Caspi, A., Poulton, R., Meier, M.H., Houts, R., Harrington, H. *et al.* Specificity of childhood psychotic symptoms for predicting schizophrenia by 38 years of age: a birth cohort study. *Psychol Med* 2013; **43**: 2077-2086.

11. Jones, H.J., Stergiakouli, E., Tansey, K.E., Hubbard, L., Heron, J., Cannon, M. *et al.* Phenotypic Manifestation of Genetic Risk for Schizophrenia During Adolescence in the General Population. *JAMA Psychiatry* 2016; **73**: 221-228.

12. Sieradzka, D., Power, R.A., Freeman, D., Cardno, A.G., McGuire, P., Plomin, R. *et al.* Are Genetic Risk Factors for Psychosis Also Associated with Dimension-Specific Psychotic Experiences in Adolescence? *Plos One* 2014; **9**.

13. Jones, H.J., Heron, J., Hammerton, G., Stochl, J., Jones, P.B., Cannon, M. *et al.* Investigating the genetic architecture of general and specific psychopathology in adolescence. *Transl Psychiat* 2018; **8**.

14. McGrath, J.J., Saha, S., Al-Hamzawi, A.O., Alonso, J., Andrade, L., Borges, G. *et al.* Age of Onset and Lifetime Projected Risk of Psychotic Experiences: Cross-National Data From the World Mental Health Survey. 2016; **42**: 933-941.

15. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015; **12**: e1001779.

16. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K. *et al.* Genome-wide genetic data on ~500,000 UK Biobank participants. *bioRxiv* 2017.

17. McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A.R., Teumer, A. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* 2016; **48**: 1279-1283.

18. Huang, J., Howie, B., McCarthy, S., Memari, Y., Walter, K., Min, J.L. *et al.* Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. *Nat Commun* 2015; **6**: 8111.

19. Smith, D.J., Escott-Price, V., Davies, G., Bailey, M.E., Colodro-Conde, L., Ward, J. *et al.* Genome-wide analysis of over 106 000 individuals identifies 9 neuroticism-associated loci. *Mol Psychiatry* 2016; **21**: 1644.

20. Maechler, M., Rousseeuw, P., Croux, C., Todorov, V., Ruckstuhl, A., Salibian-Barrera, M. *et al.* robustbase: Basic Robust Statistics R package version 0.93-2. (2018).

21. Hubert, M., Rousseeuw, P.J. & Verdonck, T. A Deterministic Algorithm for Robust Location and Scatter. *J Comput Graph Stat* 2012; **21**: 618-637.

22. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007; **39**: 906-913.

23. Band, G. & Marchini, J. BGEN: a binary file format for imputed genotype and haplotype data. 2018.

24. Tucker, G., Price, A.L. & Berger, B. Improving the power of GWAS and avoiding confounding from population stratification with PC-Select. *Genetics* 2014; **197**: 1045-1049.

25. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M. & Lee, J.J. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 2015; **4**: 7.

26. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat Communications* 2017; **8**: 1826.

27. Bulik-Sullivan, B.K., Loh, P.R., Finucane, H.K., Ripke, S., Yang, J., Patterson, N. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015; **47**: 291-+.

28. Boyd, A., Golding, J., Macleod, J., Lawlor, D.A., Fraser, A., Henderson, J. *et al.* Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* 2013; **42**: 111-127.

29. Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G. *et al.* Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* 2013; **42**: 97-110.

30. Zammit, S., Kounali, D., Cannon, M., David, A.S., Gunnell, D., Heron, J. *et al.* Psychotic Experiences and Psychotic Disorders at Age 18 in Relation to Psychotic Experiences at Age 12 in a Longitudinal Population-Based Cohort Study. *Am J Psychiat* 2013; **170**: 742-750.

31. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421-427.

32. Pardinas, A.F., Holmans, P., Pocklington, A.J., Escott-Price, V., Ripke, S., Carrera, N. *et al.* Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet* 2018.

33. Stahl, E., Forstner, A., McQuillin, A., Ripke, S., Bipolar Disorder Working Group of the PGC, Ophoff, R. *et al.* Genomewide association study identifies 30 loci associated with bipolar disorder. *bioRxiv* 2017.

34. Wray, N.R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E.M., Abdellaoui, A. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet* 2018; **50**: 668-681.

35. Genetics of Personality Consortium, de Moor, M.H., van den Berg, S.M., Verweij, K.J., Krueger, R.F., Luciano, M. *et al.* Meta-analysis of Genome-wide Association Studies for Neuroticism, and the Polygenic Association With Major Depressive Disorder. 2015; **72**: 642-650.

36. Savage, J.E., Jansen, P.R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C.A. *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat Genet* 2018; **50**: 912-919.

37. Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.R. *et al.* An atlas of genetic correlations across human diseases and traits. 2015; **47**: 1236-1241.

38. Demontis, D., Walters, R.K., Martin, J., Mattheisen, M., Als, T.D., Agerbo, E. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet* 2019; **51**: 63-75.

39. Grove, J., Ripke, S., Als, T.D., Mattheisen, M., Walters, R., Won, H. *et al.* Common risk variants identified in autism spectrum disorder. 2017: 224774.

40. 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.

41. Kendall, K.M., Rees, E., Escott-Price, V., Einon, M., Thomas, R., Hewitt, J. *et al.* Cognitive Performance Among Carriers of Pathogenic Copy Number Variants: Analysis of 152,000 UK Biobank Subjects. *Biol Psychiatry* 2017; **82**: 103-110.

42. Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F. *et al.* PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res* 2007; **17**: 1665-1674.

43. Rees, E., Kendall, K., Pardinas, A.F., Legge, S.E., Pocklington, A., Escott-Price, V. *et al.* Analysis of Intellectual Disability Copy Number Variants for Association With Schizophrenia. *JAMA Psychiatry* 2016; **73**: 963-969.

44. Coe, B.P., Witherspoon, K., Rosenfeld, J.A., van Bon, B.W., Vulto-van Silfhout, A.T., Bosco, P. *et al.* Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat Genet* 2014; **46**: 1063-1071.

45. Zheng, J., Erzurumluoglu, A.M., Elsworth, B.L., Kemp, J.P., Howe, L., Haycock, P.C. *et al.* LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* 2017; **33**: 272-279.

46. Leussis, M.P., Madison, J.M. & Petryshen, T.L. Ankyrin 3: genetic association with bipolar disorder and relevance to disease pathophysiology. *Biol Mood Anxiety Disord* 2012; **2**: 18.

47. Bi, C., Wu, J., Jiang, T., Liu, Q., Cai, W., Yu, P. *et al.* Mutations of ANK3 identified by exome sequencing are associated with autism susceptibility. 2012; **33**: 1635-1638.

48. Atwood, B.K. & Mackie, K. CB2: a cannabinoid receptor with an identity crisis. *Br J Pharmacol* 2010; **160**: 467-479.

49. Lu, H.C. & Mackie, K. An Introduction to the Endogenous Cannabinoid System. 2016; **79**: 516-525.

50. Moore, T.H., Zammit, S., Lingford-Hughes, A., Barnes, T.R., Jones, P.B., Burke, M. *et al.* Cannabis use and risk of psychotic or affective mental health outcomes: a systematic review. *Lancet* 2007; **370**: 319-328.

51. Huang, W.J., Chen, W.W. & Zhang, X. Endocannabinoid system: Role in depression, reward and pain control (Review). *Mol Med Rep* 2016; **14**: 2899-2903.

52. Manseau, M.W. & Goff, D.C. Cannabinoids and Schizophrenia: Risks and Therapeutic Potential. *Neurotherapeutics* 2015; **12**: 816-824.

53. Kendall, K.M., Bracher-Smith, M., Fitzpatrick, H., Lynham, A., Rees, E., Escott-Price, V. *et al.* Cognitive performance and functional outcomes of carriers of pathogenic copy number variants: analysis of the UK Biobank. *Br J Psychiatry* 2019: 1-8.

54. Kendall, K.M., Rees, E., Bracher-Smith, M., Legge, S.E., Riglin, L., Zammit, S. *et al.* The role of rare copy number variants in depression. *JAMA Psych* 2019.
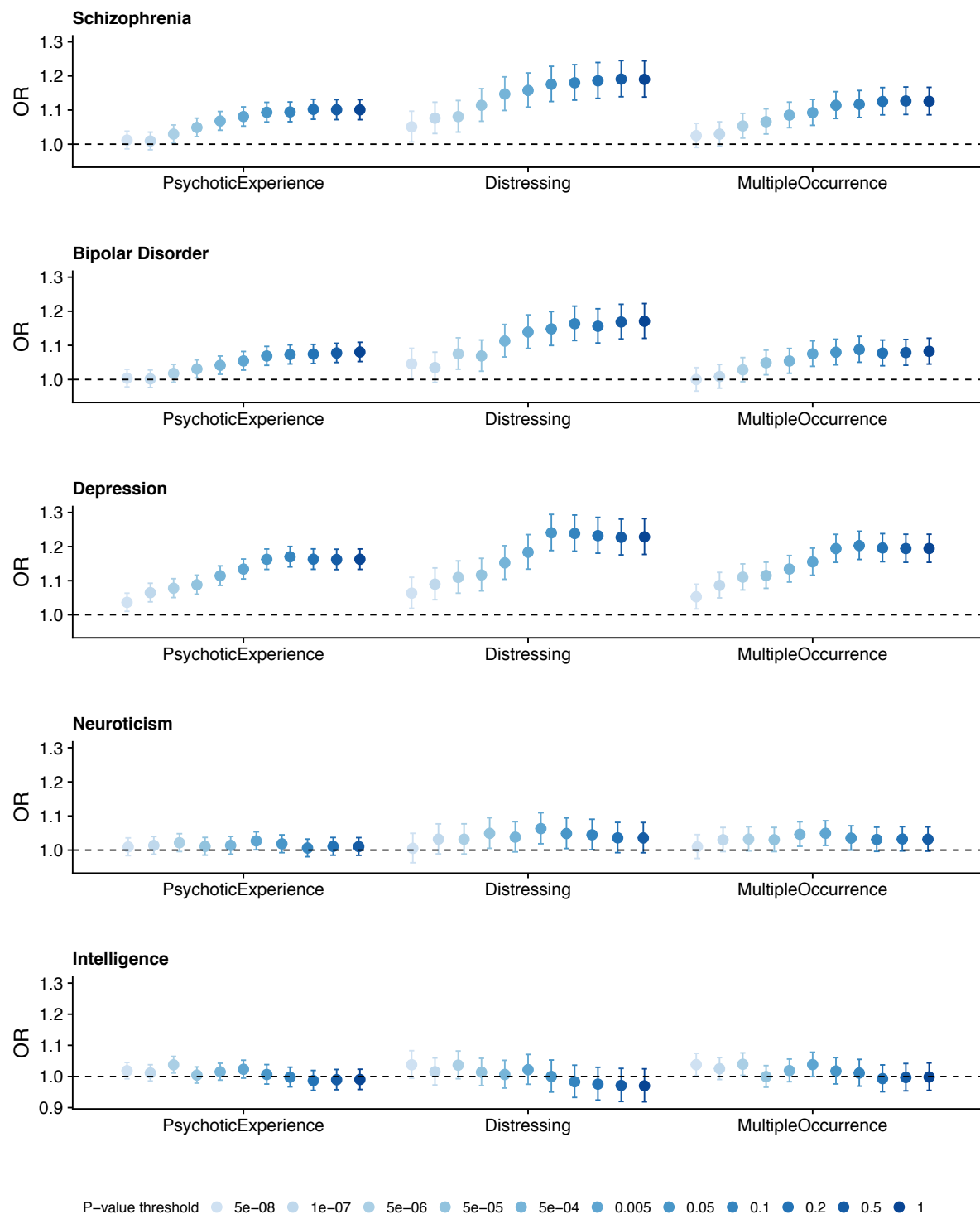
55.   Croft, J., Heron, J., Teufel, C., Cannon, M., Wolke, D., Thompson, A. *et al.* Association of Trauma Type, Age of Exposure, and Frequency in Childhood and Adolescence With Psychotic Experiences in Early Adulthood. *JAMA Psychiatry* 2018.

56.   Lahey, B.B., Krueger, R.F., Rathouz, P.J., Waldman, I.D. & Zald, D.H. A hierarchical causal taxonomy of psychopathology across the life span. *Psychol Bull* 2017; **143**: 142-186.

57.   Stochl, J., Khandaker, G.M., Lewis, G., Perez, J., Goodyer, I.M., Zammit, S. *et al.* Mood, anxiety and psychotic phenomena measure a common psychopathological factor. *Psychol Med* 2015; **45**: 1483-1493.

58.   Fry, A., Littlejohns, T.J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am J Epidemiol* 2017; **186**: 1026-1034.
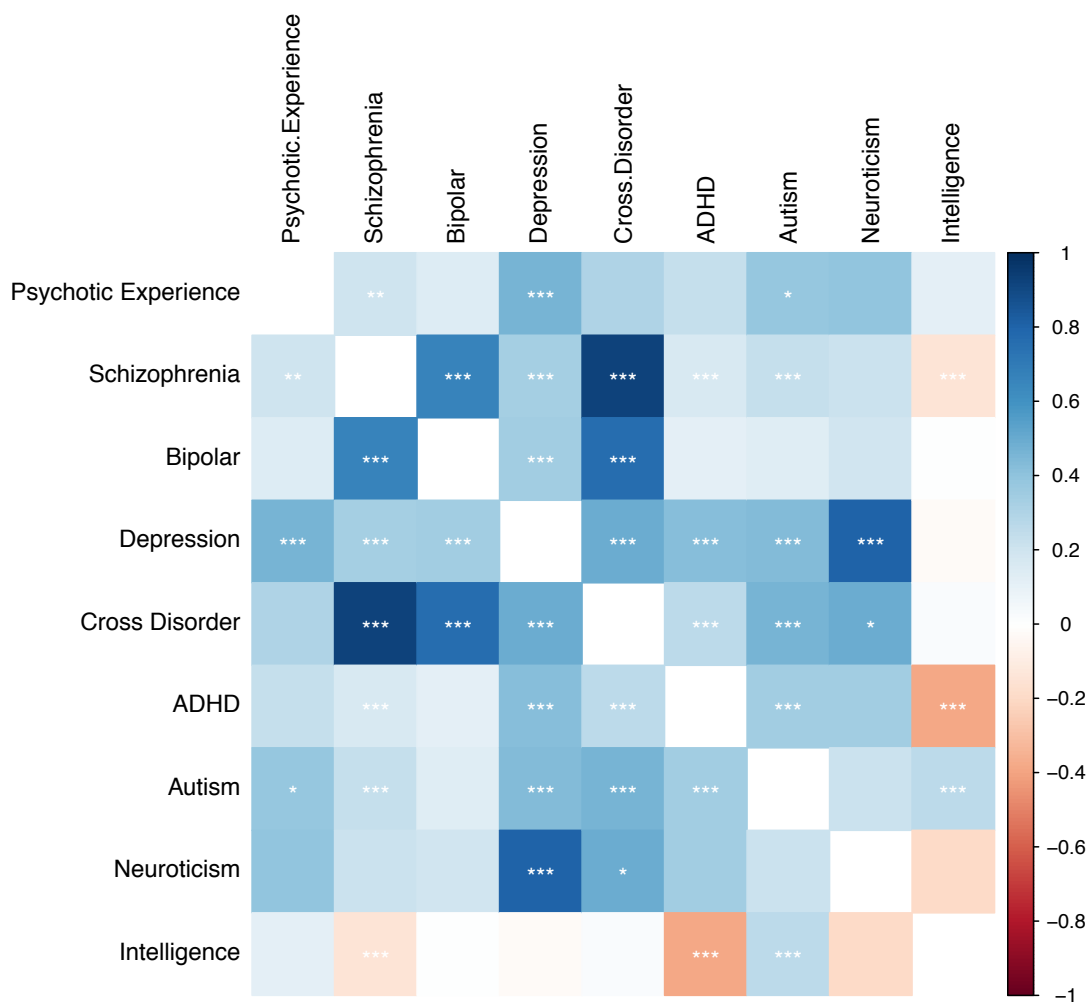
Figure 1

Manhattan plot for GWAS analyses of *any psychotic experience*, *distressing psychotic experiences* and *multiple occurrence psychotic experiences*. Dashed line represent the genome-wide significance level of $p < 5 \times 10^{-8}$ and the dotted line represents $p < 1 \times 10^{-5}$.

Figure 2

Polygenic risk score analysis. Points represent odds ratio (OR) and error bars are 95% confidence intervals. A plot with r^2 and p-values is presented in *Supplementary Figure 7.*

Figure 3

Genetic correlation analysis. Colour corresponds to the strength of the correlation (rg) and stars correspond to the statistical significance of the correlation (* P < 0.003, ** P < 0.0001, *** P < 0.00001). Positive correlations are shown in blue and negative correlations in red.

Table 1

Phenotype	SNP	CHR	BP	A1	OR (95% CI)	P	Position	Nearest gene
Any PE	rs10994278	10	62009219	T	1.16 (1.10-1.23)	3.06×10^{-8}	Intronic	*ANK3*
Any PE	rs549656827	5	35349428	G	0.61 (0.50-0.73)	3.30×10^{-8}	Intergenic	*U3, PRLR*
Distressing PE	rs75459873	1	24256312	G	0.66 (0.56-0.78)	3.78×10^{-8}	Intronic	*CNR2*
Distressing PE	rs3849810	8	53009606	A	1.22 (1.13-1.31)	4.55×10^{-8}	Intergenic	*ST18*

Genome-wide significant associations ($p < 5 \times 10^{-8}$) from GWAS analyses. Columns represent GWAS phenotype (PE = psychotic experience), variant (SNP), chromosome (CHR), base position (BP), risk allele (A1), odds ratio (OR) and 95% confidence intervals, p-value (P), position of risk variant (Position), and the nearest gene.

Table 2

Schizophrenia CNVs

	Case rate (n)	Control rate (n)	OR (95% CI)	P
Any PE	0.010 (59/5829)	0.006 (572/115807)	1.54 (1.18-2.01)	1.60×10^{-3}
Distressing PE	0.014 (28/2046)	0.006 (572/115807)	2.04 (1.39-2.98)	2.49×10^{-4}
Multiple occurrence PE	0.009 (28/3177)	0.006 (572/115807)	1.34 (0.91-1.94)	0.144

Neurodevelopmental disorder CNVs

	Case rate (n)	Control rate (n)	OR (95% CI)	P
Any PE	0.014 (83/5829)	0.009 (1058/115807)	1.54 (1.23-1.92)	1.89×10^{-4}
Distressing PE	0.017 (34/2046)	0.009 (1058/115807)	1.75 (1.24-2.48)	1.41×10^{-3}
Multiple occurrence PE	0.012 (38/3177)	0.009 (1058/115807)	1.30 (0.92-1.78)	0.139

Association of psychotic experience phenotypes with copy number variants (CNVs) previously associated with (i) schizophrenia and (ii) neurodevelopmental disorders (note all schizophrenia-associated CNVs are also included in neurodevelopmental disorders). Columns represent; phenotype (PE = psychotic experience), rate of CNVs in cases (those reporting a psychotic experience), rate of CNVs in controls (the comparator group, those who did not report a psychotic experience), odds ratio (OR) and 95% confidence intervals (CI), and p-value (P).