Characterisation of age and polarity at onset in bipolar disorder


Background

Studying phenotypic and genetic characteristics of age at onset (AAO) and polarity at onset (PAO) in bipolar disorder can provide new insights into disease pathology and facilitate the development of screening tools.

Aims

To examine the genetic architecture of AAO and PAO and their association with bipolar disorder disease characteristics.

Method

Genome-wide association studies (GWASs) and polygenic score (PGS) analyses of AAO (n = 12 977) and PAO (n = 6773) were conducted in patients with bipolar disorder from 34 cohorts and a replication sample (n = 2237). The association of onset with disease characteristics was investigated in two of these cohorts.

Results

Earlier AAO was associated with a higher probability of psychotic symptoms, suicidality, lower educational attainment, not living together and fewer episodes. Depressive onset correlated with suicidality and manic onset correlated with delusions and manic episodes. Systematic differences in AAO across cohorts and continents of origin were observed. This was also reflected in systematic differences in PAO across cohorts, continents and phenotype definitions introduce significant heterogeneity, affecting analyses.

Conclusions

AAO and PAO are associated with indicators of bipolar disorder severity. Individuals with an earlier onset show an increased polygenic liability for a broad spectrum of psychiatric traits. Systematic differences in AAO across cohorts, continents and phenotype definitions introduce significant heterogeneity, affecting analyses.

Keywords

Bipolar disorder; age at onset; polarity at onset; GWAS; polygenic score.

Copyright and usage

© The Author(s), 2021. Published by Cambridge University Press on behalf of the Royal College of Psychiatrists. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background

Bipolar disorder is highly heritable and affects approximately 1% of the population. It has a recurrent or chronic course and is associated with psychosocial impairment and reduced functioning, and it is a

* Joint first authors. † Joint last authors.
leading cause of global disease burden. Individuals usually experience their first (hypo)manic or depressive episode of bipolar disorder in adolescence or early adulthood, but often they are not diagnosed until 5 to 10 years later, especially in individuals with an earlier age at onset (AAO) or a depressive index episode. Early illness onset is associated with a more severe disease course and greater impairment across a wide range of mental and physical disorders and is a useful prognostic marker. However, pathophysiological processes leading to a disorder are thought to begin long before the first symptoms appear. Investigating the factors contributing to age and polarity (i.e. either a (hypo)manic or depressive episode) at onset could thus improve our understanding of disease pathophysiology and facilitate development of personalised screening and preventive measures. Accordingly, AAO and polarity at onset (PAO) of bipolar disorder are considered as suitable phenotypes for genetic analyses.

Genome-wide association studies (GWAS) have improved our understanding of the genetic architecture of susceptibility to bipolar disorder; however, the genetic determinants of AAO and PAO remain largely unknown. Evidence suggests that patients with an early AAO carry a stronger genetic loading for bipolar disorder risk. For example, an earlier parental AAO increases familial risk for bipolar disorder and is one of the strongest predictors of 5-year illness onset in affected offspring. Previous research has described that a higher genetic risk burden for schizophrenia may be associated with earlier AAO of bipolar disorder, but this finding did not replicate. Moreover, a recent study did not find an association of bipolar disorder polygenic score (PGS) with AAO. Thus far, GWAS for age at bipolar disorder onset have been underpowered, and a study of 8610 patients found no significant evidence for a heritable component contributing to onset age. The PAO was shown to cluster in families, but the genetic architecture of PAO has not yet been investigated.

**Aims**

To fill these knowledge gaps, we performed comprehensive analyses of AAO and PAO of bipolar disorder in the largest sample studied to date by (a) examining phenotype definitions and associations, (b) investigating whether the genetic load for neuropsychiatric disorders and traits contributes to AAO and PAO of bipolar disorder, and (c) conducting systematic GWASs.

**Method**

References to published methods are listed in Supplementary Note 1 available at https://doi.org/10.1192/bjp.2021.102.

**Study samples**

Participants with a bipolar disorder diagnosis, available genetic data and AAO information were selected from independent data-sets, including those previously submitted to the Psychiatric Genomics Consortium (PGC) Bipolar Disorder Working Group and the International Consortium on Lithium Genetics (ConLiGen). These consortia aggregate genetic data from many cohorts worldwide. Our analyses comprised 34 cohorts with 12,977 patients with bipolar disorder who have European ancestry from Europe, North America and Australia. For a description of sample ascertainment, see the Supplementary Material.

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. All procedures involving human patients were approved by the local ethics committees, and written informed consent was obtained from all patients. For details on the data-sets, including phenotype definitions and distributions, see Table 1, Fig. 1, and Supplementary Table S1.

**Definition of AAO**

The definition of age at bipolar disorder onset differed by cohort. To enhance cross-cohort comparability, we grouped the definitions into four broad categories as follows (Supplementary Table S1).

(a) Diagnostic interview: age at which the patient first experienced a (hypo)manic, mixed or major depressive episode according to a standardised diagnostic interview.

(b) Impairment/help-seeking: age at which symptoms began to cause subjective distress or impaired functioning or at which the patient first sought psychiatric treatment.

(c) Pharmacotherapy: age at first administration of medication.

(d) Mixed: a combination of the above-mentioned definitions.

Across definitions, participants younger than 8 years at onset were excluded (n = 279) because of the uncertainty about the reliability of retrospective recall of early childhood onset. The distribution of AAO was highly skewed and differed considerably between the cohorts (Table 1 and Fig. 1). Therefore, we transformed AAO in each cohort by rank-based inverse-normal transformation and used this normalised variable as the primary dependent variable in all genetic analyses. To facilitate interpretability of effect sizes, we also report results of the corresponding untransformed AAO.

**Definition of PAO**

For each cohort, PAO was defined by comparing the age at the first (hypo)manic and first depressive episode or using the polarity variable provided by the cohort. Specifically, patients were divided into three subgroups:

(a) (hypo)mania before depression (PAO-M);
(b) depression before (hypo)mania (PAO-D); and
(c) mixed (PAO-X).

The third category included patients with mixed episodes and those with a first (hypo)manic and depressive episode within the same year (Table 1). In the primary analysis, we combined patients with (hypo)mania and mixed onset and assigned this as the reference category. In secondary analyses, we excluded the patients in the mixed group.

**Phenotypic disease characteristics**

We performed phenotypic analyses of disease onset in patients with bipolar disorder type I from three cohorts: the Dutch Bipolar cohort (n = 1313) and the German PsyCourse and FOR2107 cohorts, which were analysed jointly (n = 346). We analysed the following disease characteristics, which were previously reported as being associated with disease onset and were assessed in a similar way across cohorts: lifetime delusions, lifetime hallucinations, history of suicide attempt, suicidal ideation, current smoking, educational attainment, living together with a partner, and frequency of manic and depressive episodes per year. For more detailed information, see the Supplementary Note 2 and Supplementary Table S9.

**Quality control and imputation of genotype data**

The cohorts were genotyped according to local protocols. Individual genotype data of all discovery-stage cohorts were processed with the PGC Rapid Imputation and Computational Pipeline for GWAS (RICOPILI) with the default parameters for standardised quality control, imputation and analysis. Before imputation, filters for the removal of variants included non-autosomal chromosomes, missingness ≥0.02, and a Hardy–Weinberg equilibrium test P < 1 × 10^-10.

2
<table>
<thead>
<tr>
<th>GWAS stage, dataset</th>
<th>n</th>
<th>Continent</th>
<th>Diagnosis, % bipolar disorder type I</th>
<th>Gender, % male</th>
<th>AAO, median (MAD, a range)</th>
<th>Definition of AAO</th>
<th>PAO, b n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wtccc</td>
<td>1452</td>
<td>Europe</td>
<td>89.53</td>
<td>36.85</td>
<td>24 (8.9, 9–63)</td>
<td>Impairment/help-seeking</td>
<td></td>
</tr>
<tr>
<td>tgco2</td>
<td>865</td>
<td>North America</td>
<td>100</td>
<td>33.64</td>
<td>17 (7.93, 8–46)</td>
<td>Diagnostic interview</td>
<td>PAO-M: 316 (38.92); PAO-D: 496 (61.08)</td>
</tr>
<tr>
<td>gain</td>
<td>772</td>
<td>North America</td>
<td>100</td>
<td>48.06</td>
<td>18 (7.93, 8–43)</td>
<td>Diagnostic interview</td>
<td>PAO-M: 135 (18.57); PAO-D: 440 (60.52)</td>
</tr>
<tr>
<td>stp1</td>
<td>718</td>
<td>North America</td>
<td>100</td>
<td>44.01</td>
<td>16 (7.93, 8–41)</td>
<td>Diagnostic interview</td>
<td>PAO-M: 137 (19.68); PAO-D: 420 (58.5)</td>
</tr>
<tr>
<td>gsk1</td>
<td>715</td>
<td>North America</td>
<td>89.51</td>
<td>36.36</td>
<td>17 (7.91, 8–52)</td>
<td>Diagnostic interview</td>
<td>PAO-M: 102 (14.61); PAO-D: 395 (56.59)</td>
</tr>
<tr>
<td>usc2</td>
<td>681</td>
<td>North America</td>
<td>96.18</td>
<td>48.06</td>
<td>18 (7.93, 8–45)</td>
<td>Diagnostic interview</td>
<td>PAO-M: 43 (16.33); PAO-D: 159 (50.46)</td>
</tr>
<tr>
<td>bonn</td>
<td>638</td>
<td>Europe</td>
<td>99.84</td>
<td>44.37</td>
<td>18 (7.93, 8–41)</td>
<td>Diagnostic interview</td>
<td>PAO-M: 73 (17.14); PAO-D: 238 (55.87)</td>
</tr>
<tr>
<td>ucl2</td>
<td>604</td>
<td>Europe</td>
<td>100</td>
<td>44.37</td>
<td>18 (7.93, 8–50)</td>
<td>Diagnostic interview</td>
<td>PAO-M: 54 (14.25); PAO-D: 197 (51.98)</td>
</tr>
<tr>
<td>ucl2</td>
<td>449</td>
<td>Europe</td>
<td>100</td>
<td>44.37</td>
<td>18 (7.93, 8–45)</td>
<td>Diagnostic interview</td>
<td>PAO-M: 54 (14.25); PAO-D: 197 (51.98)</td>
</tr>
<tr>
<td>Summary</td>
<td>1277</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Replication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ukw1</td>
<td>1156</td>
<td>Europe</td>
<td>75.17</td>
<td>36.15</td>
<td>23 (8.9, 8–74)</td>
<td>Impairment/help-seeking</td>
<td></td>
</tr>
<tr>
<td>dutch</td>
<td>468</td>
<td>Europe</td>
<td>100</td>
<td>42.31</td>
<td>28 (10.38, 11–63)</td>
<td>Pharmacotherapy</td>
<td>PAO-M: 316 (38.92); PAO-D: 496 (61.08)</td>
</tr>
<tr>
<td>st5</td>
<td>186</td>
<td>North America</td>
<td>100</td>
<td>53.23</td>
<td>16 (7.41, 8–51)</td>
<td>Unknown</td>
<td>PAO-M: 316 (38.92); PAO-D: 496 (61.08)</td>
</tr>
<tr>
<td>colo</td>
<td>176</td>
<td>South America</td>
<td>90.34</td>
<td>31.82</td>
<td>20 (11.86, 8–52)</td>
<td>Diagnostic interview</td>
<td>PAO-M: 316 (38.92); PAO-D: 496 (61.08)</td>
</tr>
<tr>
<td>bnmrom</td>
<td>126</td>
<td>Europe</td>
<td>100</td>
<td>42.86</td>
<td>24 (8.9, 12–56)</td>
<td>Diagnostic criteria</td>
<td>PAO-M: 316 (38.92); PAO-D: 496 (61.08)</td>
</tr>
<tr>
<td>bdters</td>
<td>125</td>
<td>Europe</td>
<td>64</td>
<td>45.6</td>
<td>28 (13.34, 8–65)</td>
<td>Impairment/help-seeking</td>
<td>PAO-M: 316 (38.92); PAO-D: 496 (61.08)</td>
</tr>
<tr>
<td>Summary</td>
<td>2237</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All data</td>
<td>15244</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

GWAS, genome-wide association study; AAO, age at onset; MAD, median absolute deviation; PAO, polarity at onset; PAO-M, mania/hypomania before depression; PAO-D, depression before mania/hypomania.

a. We calculated the median absolute deviation using 1.4826 as constant.

b. We defined three categories of polarity at onset: PAO-M, mania/hypomania before depression; PAO-D, depression before mania/hypomania, and mixed. PAO was not available for all patients. The table presents the PAO-M and PAO-D subgroups and their percentage within the individual cohorts.
Individuals were removed if they showed a genotyping rate $\leq 0.98$, absolute deviation in autosomal heterozygosity of $F_{het} \geq 0.2$, or a deviation $\geq 4$ s.d.s from the mean in any of the first eight ancestry components within each cohort. From genetic duplicates and relatives (pi-hat $>0.2$) across all samples, only the individual with more complete phenotypic information on AAO and PAO, gender and diagnosis was retained. Imputation was performed by IMPUTE2 with the Haplotype Reference Consortium reference panel.

**PGS**

We calculated PGS based on prior GWAS of attention-deficit hyperactivity disorder (ADHD), autism spectrum disorder (ASD), bipolar disorder, educational attainment (measured as years in education), major depression (MD), and schizophrenia (see Supplementary Table S3, which includes references). PGS weights were estimated with PRS-CS (see Supplement), with six scores per GWAS (with $\varphi = 1 \times 10^{-3}$, $1 \times 10^{-2}$, $1 \times 10^{-1}$, $1 \times 10^{-0.5}$, and $1 \times 10^{-0}$). We tested the associations of the PGS with the AAO and PAO by linear and logistic regressions, respectively. Gender, bipolar disorder subtype and the first eight ancestry components were included as covariates. The significance threshold was Bonferroni-corrected for 96 tests ($\alpha = 0.05/(6 \varphi$ thresholds $\times 8$ traits $\times 2$ phenotypes) = $5.2 \times 10^{-4}$).

**GWASs**

We performed a discovery GWAS on the 34 cohorts ($n = 12\,977$) and replication analyses in six additional cohorts with $n = 2\,237$ patients with bipolar disorder. As a first step, we conducted individual GWAS for each cohort with 40 or more patients using the RICOPILI workflow, using the same covariates as in the PGS analyses. Sample sizes are provided in Supplementary Tables S2 and S7. The resulting GWAS did not show an inflation of test statistics for any of the cohorts, indicating limited population stratification (Supplementary Table S2). Next, we performed a fixed-effects meta-analysis using METAL, combining the cohort-specific GWASs. For the meta-analysis summary statistics, we applied the following variant-level post-quality control parameters: imputation INFO score $>0.9$, minor allele frequency (MAF) $\geq 0.05$, and successfully imputed/genotyped in more than half of the cohorts.

The primary analyses were AAO (normalised, analysed by linear regression) and PAO (analysed by logistic regression). Secondary analyses included GWASs stratified by AAO definition and continent of origin.

We estimated the power to replicate our initial genome-wide significant finding from the discovery GWAS based on the regression coefficients using the $pwr$ package in $R$. Assuming the same effect size and MAF (beta 0.075, allele frequency 0.32) and a standardised phenotype, we had 76% power to detect the effect in our sample size of 2237 at an alpha level of 0.1. For comparison, we had 57% power to detect the effect in our discovery sample, using the more stringent genome-wide significance cut-off.

**Heritability analyses**

Next, we assessed the overall variance in AAO and PAO explained by genotyped variants (so-called single-nucleotide variant (SNV)-based heritability, $h_{SNV}^2$). For the only individual cohort with more than 1000 samples, we estimated $h_{SNV}^2$ with GCTA GREML. In this case, we validated the robustness of the $h_{SNV}^2$ estimate with
the mean of 1000 × resampling of 95% of the sample. To estimate the overall heritability of the meta-analysis summary statistics we estimated $\hat{h}^2_{GCV}$ by linkage disequilibrium score regression, for each GWAS with sample size >3000. The 95% CIs were constrained to a minimum of 0 and a maximum of 1.

## Results

### Heterogeneity of AAO and PAO across cohorts

Among the four definitions of AAO across the 34 cohorts, impairment/help-seeking was the most common in Europe and diagnostic interview the most common in North America (Table 1, Fig. 1). Across all cohorts, the median AAO was 21 years (range of medians: 16–30 years; Fig. 1). However, substantial differences in the AAO were observed between subgroups: first, the median untransformed AAO was lower in bipolar disorder type I than in type II (type I, 21 years; type II, 22 years; Kruskal-Wallis test $P = 1.8 \times 10^{-4}$, Supplementary Table S6).

Second, the AAO was lower when determined by diagnostic interview compared with other phenotype definitions (diagnostic interview, 19 years; impairment/help-seeking, 23 years; pharmacotherapy, 30 years; mixed, 22 years; $P = 2.96 \times 10^{-19}$). Third, the age was lower in North America compared with Europe (Europe, 24 years; North America, 18 years; and Australia, 19.5 years; $P = 2.0 \times 10^{-28}$). These differences across continents remained significant when including onset definitions and bipolar disorder subtype in a multivariable regression model, indicating that they are likely partially independent from the assessment strategy (Supplementary Table S6).

The majority of patients reported a depression-first PAO. Patients with depression-first were less frequent in the impairment/help-seeking than in the diagnostic interview category (55% and 59%, respectively; $P = 4.5 \times 10^{-4}$, Supplementary Fig. S1), but their proportions were similar between Europe and North America (57% and 59%, respectively; $P = 0.17$ test of proportion).

### Analyses of disease characteristics

In a meta-analysis of the Dutch and German samples, earlier AAO was significantly associated with a higher probability of lifetime delusions, hallucinations, suicide attempts, suicidal ideation, lower educational attainment and not living together (Table 2, Supplementary Tables S4 and S5). A later AAO was positively significantly correlated with a higher number of manic and depressive episodes per year (see Tables 3, and the Supplementary Note 2). Moreover, a (hypo)manic onset was significantly associated with a greater likelihood of delusions and more manic episodes per year, whereas a depressive onset was associated with a higher probability of suicidal ideation and lifetime suicide attempts.

### Associations of PGSSs with AAO and PAO

Next, we conducted analyses to evaluate whether the genetic liability for five psychiatric disorders and educational attainment were associated with the age at disease onset (Fig. 2(a) and (b) and Supplementary Table S8). After correcting for 96 tests, higher PGSSs for ASD ($\beta = -0.34$ per year s.d. increase in PGSS, s.e. $= 0.08$, $P = 9.85 \times 10^{-8}$), major depression ($\beta = -0.34$, s.e. $= 0.08$, $P = 1.40 \times 10^{-7}$), schizophrenia ($\beta = -0.39$, s.e. $= 0.08$, $P = 2.91 \times 10^{-8}$) and educational attainment ($\beta = -0.31$, s.e. $= 0.08$, $P = 5.58 \times 10^{-8}$) were significantly associated with an earlier age at bipolar disorder onset. This was not the case for ADHD or bipolar disorder PGSS. No PGSS was significantly associated with PAO (Supplementary Fig. S4, Supplementary Table S8).

### GWASs

Next, we attempted to identify individual genetic loci associated with the AAO or PAO. In our discovery GWAS using 34 cohorts, one locus was significantly associated with AAO (rs1610275 on chromosome 16; minor allele G frequency $= 0.319$, $\hat{\beta} = 0.075$ (s.e. $= 0.014$), $P = 3.39 \times 10^{-8}$, Fig. 2(c), Supplementary Table S7, Supplementary Fig. S2). This SNV mapped to an intron of the brain-expressed gene FTO (alpha-ketoglutarate dependent dioxygenase, Fig. 2(d)).
Fig. 2 Results from the genome-wide association study (GWAS), polygenic score (PGS) analyses, and heritability analyses.

(a) and (b) Results from analyses of PGS. For detailed results, see Supplementary Table S8. Significance levels: n.s., not significant, P > 0.05; nominal: P < 0.05; Bonferroni, below the Bonferroni-corrected significance threshold corrected for 96 tests (P < 5.2 × 10⁻⁴). (a) Associations of PGSs with the AAO. For interpretability, the plot shows the untransformed AAO. Significance levels are based on the analyses of the AAO after rank-based inverse-normal transformation (which was performed because the distribution of AAO was highly skewed and differed greatly across the study cohorts). (b) Associations of the top versus bottom AAO quartiles with the top versus bottom PGS quartiles. A higher odds ratio (OR) indicates an association with higher AAO.

(c) Manhattan plot of the discovery-stage AAO GWAS. (d) Locus-specific Manhattan plot of the top-associated AAO variant. (e) Estimation of the variance in different phenotype definitions explained by genotyped single-nucleotide variants (SNV) (h²SNV). For the cohort wtccc, we directly estimated h²SNV from genotype data in GCTA GREML; we estimated all other heritabilities from GWAS summary statistics using LDSC. The plot shows h²SNV estimates and s.e.

ADHD, attention-deficit/hyperactivity disorder; ASD, autism spectrum disorder; BD, bipolar disorder; cM, centi Morgan; Mbp, mega base pairs; MD, major depression; EA, educational attainment; SNV, single-nucleotide variant; cont, continent; disorder type; PAO, polarity at onset; PAO-M, mania/hypomania before depression; PAO-D, depression before mania/hypomania; PAO-X, mixed; SZ, schizophrenia.
However, this association was not replicated in an independent sample of six cohorts (Supplementary Table S7, Supplementary Fig. S2). In the replication sample \((n = 2237)\), we had 76% power to replicate this SNV at a \(P\)-value threshold of 0.1. The GWAS of PAO did not yield any genome-wide significant findings, in either primary (PAO-M-X versus PAO-D) or secondary (PAO-M versus PAO-D) analyses (Supplementary Fig. S3).

We also calculated PGs for AAO and PAO using leave-one-out summary statistics from these GWASs. The AAO PGs was nominally significantly associated with AAO \((\beta = 0.23\) years, \(s.e. = 0.08, P = 0.0087, \phi = 0.1, \text{Fig. 2(a) and 2(b)}\)) for five of six tested \(\phi\) parameters but did not withstand correction for multiple testing (Supplementary Table S8). The PAO PGs was not associated with the PAO (Supplementary Fig. S4).

### SNV-based heritability of the investigated phenotypes

We estimated the SNV-based heritability \(h^2_{\text{SNV}}\) directly from genotype data using GCTA in the only cohort large enough for this analysis, \(wtccc\). For the AAO, the \(h^2_{\text{SNV}}\) in \(wtccc\) was estimated at 0.63 \((P = 0.0026)\) (Fig. 2(e)). We evaluated the robustness of this estimate by resampling (mean \(h^2_{\text{SNV}}\) = 0.62, resampling 95% CI 0.15–1.00).

We next estimated \(h^2_{\text{SNV}}\) by linkage disequilibrium score regression (LDSC) from the GWAS summary statistics generated in the present study (Fig. 2(e)). We observed that the heritability decreased when cohorts, phenotype definitions and continents were combined (for example ‘diagnostic interview’ in North America: AAO \(h^2_{\text{SNV}} = 0.16, 95\%\ CI 0.04–0.40, \text{‘impairment/help-seeking’} in Europe: \(h^2_{\text{SNV}} = 0.03, 95\%\ CI 0–0.25, \text{all combined}\) \(h^2_{\text{SNV}} = 0.05, 95\%\ CI 0–0.12\)). As a result of the insufficient sample size, we could not estimate the \(h^2_{\text{SNV}}\) of ‘impairment/help-seeking’ in North America and diagnostic interview in Europe. For depression versus (hypo) manic and mixed PAO, \(h^2_{\text{SNV}}\) was 0.17 (95% CI 0.05–0.29) on the observed scale.

### Discussion

In our study of bipolar disorder disease onset, we first evaluated the association between AAO or PAO with several clinical indicators of severity in a sample of 1659 patients. We showed that an earlier onset is associated with increased severity, demonstrating and replicating the clinical relevance of these phenotypes. Next, we performed genetic analyses including 12 977 patients from 34 cohorts. Here, we demonstrated that higher genetic risk for ASD, major depression, schizophrenia and educational attainment is associated with an earlier AAO, providing evidence that the age at bipolar disorder onset is influenced by a broad liability for psychiatric illness.

Third, we performed GWAS to identify genetic variants associated with the AAO and PAO, which did not yield any replicated associations. Fourth, we outlined the extent to which age (and, partly, polarity) at onset varies across cohorts, depending both on the continent of recruitment and on the diagnostic instrument used to determine the AAO.

Finally, we showed that this substantial phenotypic heterogeneity affects the heritability of the phenotype, which decreased when multiple cohorts with different diagnostic instruments were combined. This analysis emphasises how genetic analyses are hampered by phenotypic heterogeneity.

### Illness onset is associated with disease course

In a first set of analyses, we confirmed the clinical relevance of disease onset phenotypes in bipolar disorder. Age at bipolar disorder onset was associated with important illness severity indicators, such as suicidality, psychotic symptoms and lower educational attainment, thereby replicating findings of previous studies.\(^{12,27}\) Furthermore, patients with a depressive bipolar disorder onset had an increased reported lifetime suicidality, whereas those with a (hypo) manic onset were more likely to experience delusions and more manic episodes per illness year. Contrary to previous evidence in a US (but not in a French) sample, we observed that an earlier onset was associated with fewer episodes per illness year.\(^{26}\) Of note, when not normalising for the illness duration, the AAO was, as expected, positively correlated with the number of episodes (see Supplementary Note 2).

### Increased genetic scores for neuropsychiatric phenotypes predict an earlier illness onset

Higher PGs for schizophrenia, major depression, ASD and educational attainment were significantly associated with a lower AAO, and none of the tested PGs were significantly associated with PAO. Our findings support the hypothesis that a general liability for psychiatric disorders influences an earlier age of onset in bipolar disorder. Alternatively, an earlier onset may also reflect the broader phenotypic spectrum sometimes captured in early-onset bipolar disorder. Unexpectedly, and in contrast to several other disorders (for example multiple sclerosis), where the strongest genetic risk factors for disease liability are also the most important genetic factors associated with an earlier disease onset,\(^{6,27}\) we did not find a significant association between bipolar disorder PGs and the age at bipolar disorder onset. Statistical power may have influenced this result, as the sample sizes of both the schizophrenia and major depression GWASs were larger than that of the bipolar disorder GWAS, improving the predictive ability of these PGs compared with the bipolar disorder PGs.

The described significant relationship of higher educational attainment PG with an earlier AAO may seem counterintuitive. However, several studies described a significant association, genetic correlation and causal relationship between a higher educational attainment and bipolar disorder risk.\(^{28,29}\) Our findings demonstrate that a high educational attainment PG is not only a risk factor for bipolar disorder but also associated with an earlier onset of the disorder.

### Lack of replication of the GWAS finding

We have conducted two GWASs to identify individual loci influencing the age and polarity at bipolar disorder onset, possibly independently of affecting lifetime disorder risk. Our discovery GWAS prioritised a genome-wide significant locus associated with the AAO. However, the lack of replication suggests that this finding may have been false-positive. This failure to replicate could have been because of insufficient statistical power in the replication sample, as our power analysis did not account for the likely phenotypic and genetic heterogeneity across cohorts and may thus have underestimated the necessary sample size. Importantly, the replication sample was more ethnically diverse than the discovery sample, which reduced the statistical power. The PAO GWAS, with its lower sample size and dichotomous phenotype, did not identify any genome-wide significant locus.

We also calculated an AAO PG using our GWAS and tested it on our sample. Although the effect size of this PG on the AAO was substantial (0.23 years per unit change in the PG), the association was only nominally significant.

### The heterogeneity of phenotype definitions

A striking finding of our study was the systematic difference in the AAO distribution across cohorts, continents and assessment strategies. Although the assessment strategies varied considerably by continent, with diagnostic interview being mainly used in North...
Phenotypic heterogeneity affects genetic analyses

Interestingly, the systematic differences in AAO phenotypes across cohorts are reflected in heritability estimates: we observed the highest SNV-based heritability $h_{SNV}^2$ when onset was established by diagnostic interview and the lowest when it was captured with more health system-specific and subjective measurements, such as item 4 of the Operational Criteria Checklist for Psychotic Illness (impairment/help-seeking). Moreover, $h_{SNV}^2$ estimates approached zero when all samples were combined in our primary analysis ($h_{SNV}^2 = 0.05; 95\% CI 0.0–0.12$), underscoring the strong impact of phenotypic heterogeneity. For PAO-M/X versus PAO-D, we observed significant $h_{SNV}^2$ estimates, demonstrating that genetic factors contribute to the polarity at bipolar disorder onset.

Thus, we not only showed systematic heterogeneity in a clinically relevant psychiatric phenotype across cohorts but also provided direct evidence for how this heterogeneity can hamper genetic studies. Similarly, a recent investigation demonstrated that the phenotyping method (for example diagnostic interview versus self-report) significantly influenced heritability estimates, GWAS results and PGS performance in analyses of major depression susceptibility, with broader phenotype definitions resulting in lower heritability estimates. These results indicate that although increasing sample sizes generally improves the power to detect significant associations, larger samples are no silver bullet: careful phenotype harmonisation and uniform recruitment strategies are likely at least as important.

Limitations

In addition to diverse phenotype definitions originating from different ascertainment methods, as described above, several factors may have limited the cross-cohort comparability of the AAO and PAO. These factors include differences in the definition and ascertainment of the age at bipolar disorder onset and in how bipolar disorder was diagnosed across cohorts and continents. Such differences can lead to bias, affecting genetic analyses. For example, as patients diagnosed with bipolar disorder type II show, on average, later ages at onset than patients with bipolar disorder type I, differing proportions of bipolar disorder subtypes across cohorts may have an impact on AAO analyses. Therefore, we included the bipolar disorder subtype as a covariate in our genetic analyses to control for this confounder. Still, this cross-cohort heterogeneity has likely reduced our statistical power.

Given that, for all included cohorts, the disease onset phenotypes were assessed retrospectively, measurement errors associated with interrater reliabilities and recall bias may have occurred across cohorts. For example, hypomania was likely underreported, potentially biasing the PAO towards depression. Notably, such potential issues are not specific to the present study but may affect all retrospective analyses of psychiatric phenotypes. Nevertheless, differences in the diagnosis of bipolar disorder and the ascertained phenotypes between cohorts might have exacerbated these problems. Therefore, future studies should focus on compiling clinically more homogeneous, phenotypically better-harmonised data-sets instead of only assembling the largest possible sample.

Furthermore, the rank-based inverse normal transformation of the AAO phenotype may have affected the GWAS and heritability analyses. We conducted this transformation because, first, the original AAO distribution was highly skewed and thus not suitable for linear regression and, second, the AAO differed significantly between cohorts, which could have biased the meta-analysis. However, by transforming the data, only the rank and not the absolute differences in onset between patients was maintained, reducing the interpretability of the phenotype and the genetic effects.

We performed both SNV-level and polygenic score associations using a structured meta-analysis, which mitigates some of the noise introduced by phenotypic heterogeneity. However, we were unable to account for differences in the underlying genetic aetiology of the phenotypes across cohorts. As described above, phenotypic heterogeneity is an important limitation of our study and should be considered in future phenotype and genetic analyses. Our results need to be interpreted in light of these limitations.

Implications

Phenotypes of bipolar disorder onset are clinically important trait measures contributing to the well-known clinical and biological heterogeneity of this severe psychiatric disorder. Genetic analysis of AAO and PAO may lead to a better understanding of the biological risk factors underlying mental illness and support clinical assessment and prediction. Our study provides evidence of a genetic contribution to age and polarity at bipolar disorder onset but also demonstrates the need for systematic harmonisation of clinical data on bipolar disorder onset in future studies.

Janos L. Kalman, MD, Institute of Psychiatric Phenomics and Genomics (IPPG), University Hospital, LMU Munich, Germany; Department of Psychiatry and Psychotherapy, University Hospital Munich, Munich, Germany, and International Max Planck Research School for Translational Psychiatry, Germany; Loes G. Oudesluys, PhD, Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, USA; Annabel Vreeker, PhD, Department of Child and Adolescent Psychiatry Psychology, Erasmus University Medical Centre, Sophia Children’s Hospital, the Netherlands; Andrew McQuillin, PhD, Division of Psychiatry, University College London, UK; Ella A. Stahl, PhD, Division of Psychiatric Genomics, Mount Sinai School of Medicine, USA; Douglas Ruderfer, PhD, Division of Genetic Medicine, Department of Medicine, Vanderbilt Genetics Institute, Vanderbilt University Medical Center, USA; Department of Biomedical Informatics, Vanderbilt University Medical Center, USA, and Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center, USA; Maria Grigorou-Serbanescu, PhD, Alexandru Obreja Clinical Psychiatric Hospital, Bucharest, Romania; Georgia Panagiotaropoulou, MSc, Department of Psychiatry and Psychotherapy, Charles-Universitätsmedizin, Germany; Stephan Ripke, MD, PhD, Analytic and Translational Genetics Unit, Massachusetts General Hospital, USA; and Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, USA; Tim B. Bigdell, PhD, Department of Psychiatry and Behavioral Sciences, SUNY Downstate Health Sciences University, USA; and VA NY Harbor Healthcare System, USA; Frederike Stein, MD, Department of Psychiatry and Psychotherapy, Philippus-Universität Marburg, Germany; Tina Meller, Dr rer nat, MSc, Department of Psychiatry and Psychotherapy, Philippus-Universität Marburg, Germany, and Center for Mind, Brain and Behavior (CMBB), Germany; Susanne Meierer, MSc, Institute for Translational Psychiatry, Westfälische Wilhelms-Universität Münster, Germany; and Institute for Translational Neurosciences, University of Münster, Germany; Helena Petlin, MSc, International Max Planck Research School for Translational Psychiatry, Germany, and Max Planck Institute of Psychiatry, Germany; Fabian Street, PhD, Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Germany; Sergi Papiol, PhD, Institute of Psychiatric Phenomics and Genomics (IPPG), University Hospital, LMU Munich, Germany; Department of Psychiatry and Psychotherapy, University Hospital Munich, Munich, Germany, and Centre de Investigacion Biomédica en Red de Salud Mental (CIBERSAM), Spain; Mark J. Adams, PhD, Division of Psychiatry, University of Edinburg, UK; Rolf Adolfsson, MD, PhD, Department of Clinical Sciences, Medical Faculty, Umeå University, Sweden; Kristina Adourian, MD, PhD, Institute of Psychiatric Phenomics and Genomics (IPPG), University Hospital, LMU Munich, Germany, and Department of Psychiatry and Psychotherapy, University Hospital Munich, Munich, Germany; Ingrid Agartz, MD, PhD, Department of Clinical Neuroscience, Centre for Psychiatry Research, Karolinska Institutet, Sweden; Department of Psychiatric Research, DIakonhjemmet Hospital, Norway, and NORMENT Centre, Division of Mental Health and
Kalman et al.

Acknowledgments

The authors thank Jackie Kleing, Board-certified Editor in the Life Sciences SL3, for editing assistance with the manuscript. AGMA-Australia sample: we thank Gin Mafii, Colleen Loo, and Michael Breekap for their contribution to clinical classifications of a subset of patients and also Andrew Frankland for his work in collecting clinical record data. WTCCC sampler study makes use of data generated by the Wellcome Trust Case Control Consortium. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.uk.org.uk. psychiatrists: we thank those who participated in the clinical assessment of patients in France (C. Henry, S. Gard, J.P. Kahn, L. Zanouy, R.F. Cohen and O. Wajsbrot-Elgrabli) and thank the patients for their participation.

Author contributions

Concept and design: Janos L. Kalman, Loes M. Olde Lohuis, Annabel Vreeker, Till F.M. Andlauer, Thomas G. Schulte and Roel A. Ophoff. Analysis and interpretation of data: Janos L. Kalman, Loes M. Olde Lohuis, Annabel Vreeker and Till F.M. Andlauer. Drafting of the manuscript: Janos L. Kalman, Loes M. Olde Lohuis, Annabel Vreeker and Till F.M. Andlauer. Supervision, and critical revision of the manuscript: Thomas G. Schulte, Roel A. Ophoff, Francis J. McMahon, Jordan W. Smoller and Martin Alda. All other authors provided data, contributed ideas and suggestions for analyses, interpreted results and revised the final manuscript.

Funding

The funding details can be found in the Supplementary File 1.

Declaration of interest


Supplementary material

View supplementary material for this article, please visit https://doi.org/10.1192/bjp.2021.102.

Data availability

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

References
