Early onset depression: characterising developmental trajectories and the role of neuropsychiatric genetic risk variants

Rice, F., Riglin, L., Thapar, A.K., Heron, J., Anney, R., O’Donovan, M.C., Thapar, A.

Key points

Question: Do neuropsychiatric disorder genetic risk variants influence developmental trajectories of depression in youth?

Findings: Distinct depression trajectory classes were identified. A late-adolescent-onset class (17.3% of the sample) showed a typical depression trajectory and was associated with major depressive disorder risk alleles. An early-adolescent-onset class (9.0%) showed clinically significant symptomatology at age 12 and was associated with neurodevelopmental, schizophrenia and ADHD, risk alleles and childhood neurodevelopmental traits.

Meaning: Depression in youth is highly heterogeneous. Findings are consistent with emerging evidence for a neurodevelopmental component to some cases of depression and that this is more likely when onset is very early.
Abstract

Importance
Depression often first manifests in adolescence. Thereafter individual trajectories vary substantially but it is not known what shapes depression trajectories in youth. Adult studies suggest that genetic risk for schizophrenia, a psychiatric disorder with a neurodevelopmental component, may contribute to earlier onset depression.

Objective
To test the hypothesis that there are distinct trajectories of depressive symptoms and that genetic liability for neurodevelopmental psychiatric disorders (schizophrenia, ADHD), as well as for Major Depressive Disorder (MDD), contribute to early-onset depression.

Design, setting and participants
The ALSPAC (Avon Longitudinal Study of Parents and Children) study is an ongoing prospective longitudinal population-based cohort that has been collecting data since September 1990 including 7543 adolescents with data on depressive symptoms at multiple time points.

Main outcome measures
Trajectories based on self-reported depressive symptoms dichotomised by the clinical cut-point. MDD, schizophrenia and Attention Deficit Hyperactivity Disorder (ADHD) polygenic risk score (PRS) were predictors.

Results
In 7543 adolescents with depression data on more than one assessment point between age 10.5 years (mean age 10.64, SD=.25) and 18.5 years (3568 male; 3975 female), three trajectory classes were identified: persistently low (73.7%), late-adolescent-onset (17.3%), and early-adolescent-onset (9.0%). The late-adolescent-onset class was associated with MDD genetic risk only (OR MDD_PRS=1.27, 95% CI, 1.09-1.48, p=.003). The early-adolescent onset class was also associated with MDD genetic risk (OR MDD_PRS=1.24, 95% CI, 1.06-1.46, p=.007) but additionally with genetic risk for neurodevelopmental disorders (OR schizophrenia_PRS
OR ADHDPRS = 1.32, 95% CI, 1.13-1.54, p<.001) and childhood ADHD and neurodevelopmental traits.

Conclusions and relevance

We found evidence of distinct depressive trajectories, primarily distinguished by age-at-onset. The more typical depression trajectory with onset of clinically significant symptomatology at age 16 was associated with MDD genetic risk. The less common depression trajectory, with a very early onset, was particularly associated with ADHD and schizophrenia genetic risk and, phenotypically, with childhood ADHD and neurodevelopmental traits. Findings are consistent with emerging evidence for a neurodevelopmental component to some cases of depression and suggest this is more likely when onset is very early.

Keywords: neurodevelopmental, ADHD, depression, schizophrenia, longitudinal, trajectory, ALSPAC, genetic, polygenic
Introduction

Major depressive disorder (MDD) is the most common mental disorder and a leading cause of disability\(^1\), even subthreshold depressive symptoms are associated with functional impairment and future mental health problems\(^2,3\). Depression often first manifests in adolescence\(^4-6\) and thereafter, individual trajectories of depressive symptomatology vary substantially\(^7\). A family history of depression and an early age-of-onset are each associated with a more chronic symptom course in adults with MDD\(^8-10\) but it is not known what shapes early depression trajectories in youth.

Depression has a complex multifactorial etiology including a moderate heritable component\(^4,11,12\). Longitudinal and family studies show strong continuity between both adolescent-onset depressive disorder and symptoms with depression in adult life, but there are also developmental differences between depression in children, adolescents and adults\(^4\). For instance, clinical follow-up studies of very early-onset depression (average age-of-onset=10.7 years) report high rates of heterotypic continuity, where, depression is often followed by a different type of clinical disorder\(^13-15\). Twin studies also show differences in the genetic etiology of very early-onset depressive symptoms compared to those arising in mid to late adolescence\(^16-18\). At the molecular level, a recent genome-wide association study (GWAS) of adults with MDD found evidence of differences in the genetic architecture of depression where a relatively early age-of-onset (before the median age-of-onset of 27 years) was associated with genetic liability to schizophrenia, an association not seen for later-onset depression which was instead associated with MDD risk alleles\(^19\). Similar findings have been reported for emotional problems (symptoms of depression and anxiety) in that emotional problems in childhood were associated with schizophrenia risk alleles but in adult life they were additionally associated with MDD genetic risk\(^20\). The association of schizophrenia risk alleles with childhood emotional problems was particularly pronounced in those with emotional problems in both childhood and adulthood suggesting that persistent emotional symptoms beginning early may drive the association with schizophrenia risk.
alleles. As schizophrenia genetic risk is thought to involve an early neurodevelopmental component\textsuperscript{21,22}, the role of genetic risk for other neurodevelopmental disorders in early-onset depression may be important to consider. In particular, genetic risk for ADHD, a common childhood-onset neurodevelopmental disorder, may be important in early-onset depression because cross-sectional and longitudinal cohort studies show heightened rates of depression in children with ADHD\textsuperscript{23-25} which may be partly due to the strong genetic correlation between ADHD and depression ($r_g = .424$)\textsuperscript{26,27}.

Here we test the contribution of neuropsychiatric disorder genetic risk variants, specifically genetic liability to MDD, schizophrenia and ADHD, to early depression trajectories. Schizophrenia and ADHD were selected in addition to MDD as they show moderate to high genetic correlations with major depression\textsuperscript{27}, there is evidence linking schizophrenia PRS to early-onset depression\textsuperscript{19,20} and epidemiological and clinical evidence\textsuperscript{15,23-25} that ADHD may be an important antecedent of depression. Estimates of genetic liability to the disorders in the form of polygenic risk scores were derived from risk alleles defined in the largest available GWAS of those disorders. We did not have a specific hypothesis for bipolar disorder genetic risk because existing studies reporting conflicting results about the phenotypic relationship between early-onset depression and bipolar disorder\textsuperscript{13,15,28} with little evidence that this is stronger for early-onset depression. Bipolar disorder also differs from ADHD and schizophrenia in that evidence suggests it is less neurodevelopmental in origin\textsuperscript{21,22}. However, for completeness we included bipolar polygenic risk scores in our analyses (eTable2a). We hypothesised that ADHD and schizophrenia genetic risk would show association with early-onset depression. We hypothesised that depression genetic risk would be associated with depression with an onset later in adolescence.

**Methods**

The Avon Longitudinal Study of Parents and Children (ALSPAC) is an ongoing population-based prospective longitudinal UK birth cohort\textsuperscript{29,30}. Data collection began in September 6\textsuperscript{th},
The enrolled core sample consisted of 14,541 pregnant women living in Avon, England, with expected delivery dates between April 1, 1991, and December 31, 1992. Of these, 13,988 children were alive at 1 year. An additional 713 children who would have been eligible but whose mothers did not enrol during pregnancy were enrolled after age 7 years giving a total sample of 14,701 alive at 1 year. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. All participants provided written informed consent. The study website contains details of all the data that is available through a fully searchable data dictionary (http://www.bristol.ac.uk/alspac/researchers/access/). For families with multiple births, we included the oldest sibling. Individuals were included in analyses when data on the primary outcome of depressive symptoms were available for at least two time points (N=7543). Numbers of individuals with data available at different time points are in Figure 1 (Supplement).

Depressive symptoms were reported by the young person at six time points (ages 10.5, 12.5, 13.5, 16.5, 17.5 and 18.5 years) on the short Mood and Feelings Questionnaire (sMFQ). This is a well-validated symptom checklist\textsuperscript{31-33} which includes 13 items about mood symptoms during the past 2-weeks (rated 0 (not true), 1 (sometimes true) or 2 (true); score range 0-26). Scores above 11 represent clinically significant symptoms\textsuperscript{31,33} and we analysed individuals scoring above and below this to examine trajectories of clinically significant symptomatology.

Polygenic risk scores (PRS) for MDD, schizophrenia, and ADHD were generated in study individuals as the standardized mean number of disorder risk alleles in approximate linkage equilibrium ($R^2<0.20$), weighted by genome-wide association study allele effect size, derived from data of imputed autosomal single-nucleotide polymorphisms. All packages used in this analysis used Stata 13.0 to implement the PLINK toolset (http://zzz.bwh.harvard.edu/plink/). Code is available at https://github.com/ricanney/stata. In brief, best guess genotype
underwent additional marker and individual quality control. Individuals were excluded on the basis of excessive heterozygosity (> 4 standard deviations (SD) from sample mean), relatedness (> 3 SD from sample mean) and genotype missingness (>2%). Markers were excluded if they were rare, (minor allele count less than 5), had high levels of missingness (>2%) or deviated from Hardy-Weinberg equilibrium (p≤10^{-10}) or from reference MAF (>10%) (Supplement).

Scores were derived from MDD, ADHD and schizophrenia weights for 152,536, 103,041 and 27,336 SNPs respectively. Risk alleles were defined as those associated with case status in the most recent Psychiatric Genomics Consortium analyses of MDD, ADHD and schizophrenia at a threshold of $P < .50$ for depression and ADHD and $P < .05$ for schizophrenia as these thresholds maximally capture phenotypic variance\textsuperscript{26,27,34-36}. Genome-wide association study discovery sample sizes were: 130,664 cases and 330,470 controls for MDD, 20,183 cases and 35,191 controls for ADHD, and 35,476 cases and 46,839 controls for schizophrenia. All PRS were standardized prior to analysis so odds ratios represent one standard deviation change. (eTable 2a for bipolar PRS).

Phenotypic measures of neurodevelopmental problems (DSM-IV\textsuperscript{37} diagnoses of childhood ADHD, social communication problems and pragmatic language difficulties at age 7), psychotic experiences (ages 12 and 17) family history of severe depression and schizophrenia and maternal education were used (eAppendix).

Analysis

We characterised depression trajectories of symptoms dichotomised by clinical cut-point (N=7543) using latent class growth analysis (LCGA) in Mplus version 8\textsuperscript{38}. This is a probability-based technique used to identify an optimum number of distinct patterns (classes) of growth (change) in the longitudinal depression scores of individuals\textsuperscript{39}. Models were run with increasing numbers of classes starting with a one-class solution specifying
both linear and quadratic change with 500 random starting values and 50 optimisations. Residual variances were allowed to vary across measurement points. A maximum likelihood parameter estimator for which standard errors are robust to non-normality (MLR) was used. To examine associations with categorical variables e.g. gender, the DCAT auxiliary option in MPlus was used. A bias-free three step approach in MPlus (R3STEP) estimated the associations between continuous hypothesised predictor variables (PRS) and trajectory class\textsuperscript{40,41,42}. Model selection was informed by model fit indices and interpretability as recommended\textsuperscript{43}. Full Information Maximum Likelihood (FIML) estimation was used in MPlus and included all individuals with more than one depression assessment in analyses (eTable 1). For tests of PRS association with trajectory class, we re-ran analyses using inverse probability weighting (IPW)\textsuperscript{44} to address any potential bias caused by participant dropout. The pattern of results was similar (eTable 4).

**Results**

Depression symptom trajectories

A three class trajectory model provided the best fit to the data and provided results that were most readily interpretable (eTable 1). Figure 1 shows the three distinct trajectory classes – a persistently low class (73.7%), a late-adolescent-onset class (17.3%), and an early-adolescent-onset class (9.0%). In the early-adolescent-onset class, the probability of clinically significant depression was first elevated (as indicated by a probability of clinically significant depression symptoms of .44) at age 12.5 years which rose to .52 at 13.5 years. In the late-adolescent-onset class, the probability of clinically significant depression (probability = .47) was first elevated at age 16.5 years and rose at 17.5 years (.57). Both elevated trajectories were associated with a diagnosis of MDD (assessed by the CIS-R\textsuperscript{45} at age 17.5) providing ‘validation’ of the trajectory classes (late-adolescent-onset 34.4%; early-adolescent onset 22.8%, low 1.5%, overall difference $\chi^2(2) = 193.70$, $p = .001$). The estimated proportion of females was 45.8% in the low class and was higher but did not differ between the early-adolescent (74.3%) and late-adolescent-onset classes (73.2%) (Table 2).
Neuropsychiatric PRS and trajectory class

As shown in Table 1, the late-adolescent-onset class was associated with higher MDD PRS only (OR=1.27, 95% CI=1.09, 1.48, p=.003). The early-adolescent-onset class was associated with higher ADHD, schizophrenia and MDD PRS (OR ADHD PRS=1.32, 95% CI=1.13, 1.54, p<.001; OR schizophrenia PRS=1.22, 95% CI=1.04, 1.43, p=.013; OR MDD PRS=1.24, 95% CI=1.06, 1.46, p=.007). Post-hoc, we examined the association with all three psychiatric PRS and trajectory class to examine which PRS contributed most strongly (Table 1). As expected the PRS were correlated (eTable 2b). Multivariate analysis showed that the strongest association with the early-adolescent-onset class was observed for ADHD PRS, that the association with schizophrenia PRS was retained, and that the association with MDD PRS became non-significant (Table 1). Results for the late-adolescent-onset class remained the same. Bipolar PRS was not associated with trajectory classes (eTable 2a). Including ancestry derived principal components did not alter results (eTable 3).

We tested whether the trajectory classes differed phenotypically on traits conceptually related to ADHD PRS (childhood ADHD and neurodevelopmental traits) and schizophrenia PRS (psychotic experiences). For childhood neurodevelopmental traits, there is clear evidence that this is associated with both ADHD and ADHD PRS\(^{46,47}\), for psychotic experiences, there is inconsistent evidence that this is linked with psychosis and schizophrenia PRS\(^{48,49}\). (Table 2). Individuals in the early-adolescent onset class had higher rates of childhood ADHD (6.3%) than the late-adolescent onset (0.9%) and the low classes (1.7%) and more social communication and pragmatic language problems (Table 2). Proportions scoring above the standard cut points were: 20.7% early-onset, 4.2% later-onset, 5.8% low, for social communication and 13.3% early-onset, 2.1% later-onset, 1.4% low, for pragmatic language. These differences distinguished the early-adolescent and late-adolescent onset classes (Table 2). For psychotic experiences, these distinguished the early-adolescent and late-adolescent onset classes only at age 12.
Discussion

This study identified substantial variation in the developmental trajectories of depression from childhood to early adult life, and moreover, that this is partly attributable to MDD, schizophrenia and ADHD risk alleles. We found evidence of distinct depressive trajectories, primarily distinguished by age-at-onset. We found that the more common, ‘typical’, developmental trajectory, with onset after puberty and persistence into early-adulthood was associated with elevated genetic risk for depression, indexed by MDD PRS. In contrast, we found that depressive symptoms defined by a very early-onset (by age 12) were associated with all neuropsychiatric genetic risk scores assessed, with the multivariate analysis showing that the association was strongest for ADHD PRS. Phenotypically, childhood neurodevelopmental difficulties (ADHD, pragmatic language and social communication difficulties) differentiated the depression trajectories which were elevated only in the early-adolescent onset group with rates increased by 5 to 7-fold in the early-adolescent onset group. Psychotic experiences differentiated the groups at age 12 only. This may be driven by depressive symptom differences between the groups at age 12 (Figure 1) given the reported association between psychotic experiences and depression and an inconsistent association with psychotic experiences and schizophrenia PRS. The findings are consistent with a growing body of literature showing that depression has a heterogeneous etiology partly indexed by age of onset. In particular, studies of adult MDD and of symptoms measured continuously in population-based samples illustrate that a relatively earlier onset is more strongly associated with schizophrenia polygenic risk. We find an additional contribution from ADHD polygenic risk scores. The implication of those results is that early and later adolescent onset depression differ to some extent with respect to the risk factors involved and that the earlier onset disorder is more strongly influenced by neurodevelopmental factors than depression with a more typical onset in later adolescence or early adulthood. This is consistent with a number of observations from epidemiological, family and clinical studies. First, several family and clinical follow-up studies suggest that
childhood-onset depression might differ etiologically from adolescent-onset depression\textsuperscript{52-55}. Second, the epidemiology of very early-onset depression differs from that of depression with onset in mid-to-late adolescence in the gender ratio of affected individuals and long-term psychiatric outcomes\textsuperscript{13,56}. Third, neurodevelopmental difficulties including speech abnormalities and poor motor skills are particularly associated with early-onset rather than adolescent- or adult-onset depression\textsuperscript{15,57,58}. Fourth, substantial clinical evidence shows that children with ADHD, a common neurodevelopmental disorder, are at elevated risk of subsequent depressive symptoms, suicide attempt and emotional problems when they grow up\textsuperscript{25,59-62}. Indeed, theory suggests neurodevelopmental difficulties as one route to emotional disturbance through the repeated experience of academic failure and peer rejection\textsuperscript{63} although ADHD and depression may also be associated due to common risk factors\textsuperscript{64}. A clinical issue is that the response to antidepressant medication\textsuperscript{65-68} in youth is not as good as it is in adults and evidence suggests the response to tricylics may differ in pre-pubertal versus post-pubertal depression. One possibility is that more ‘neurodevelopmental’ depression shows a different type of treatment response.

The present study indicates that genetic risk for ADHD and schizophrenia in the general population is associated with a persistent, early-onset trajectory of depressive symptoms. Such effects could operate through overlapping biological pathways as well as evocative gene-environment correlation where genetic factors influence traits which then affect environmental exposures (e.g. victimization) associated with depression. Irritability, which is common in children with ADHD and other neurodevelopmental disorders, is indexed by genetic risk for ADHD in youth\textsuperscript{69}, and has been shown to increase risk for later depression\textsuperscript{70,71} may be a potential route through which ADHD genetic risk increases the likelihood of mood problems. Among those with early-onset depression, we did not identify the equal gender ratio of affected males and females that has often been reported when depression onset is very early\textsuperscript{4,72}. This was somewhat surprising. Several factors may have contributed to this. First, some research suggests that depression is particularly likely in
females with neurodevelopmental disorders which may imply that high neurodevelopmental risk is more likely to manifest as mood disorder in females\textsuperscript{24,47,73}. Second, while it is generally accepted that self-reports of adolescent mood (as used in the present study) are reliable, children with neurodevelopmental disorders who are predominately male may under-report their mood symptoms compared to typically developing children\textsuperscript{74}. This raises the possibility that the reliance on self-reported mood necessary in the present study due to repeated longitudinal assessments (see below) may mean that some individuals at high neurodevelopmental risk may have been misclassified. Finally, polygenic risk scores alone are unlikely to be able to reliably classify children’s risk of developing different types of depression trajectories. However, collectively results converge to suggest that neurodevelopmental phenotypes (ADHD, social communication and pragmatic language difficulties) and neurodevelopmental genetic risk indicates a greater probability of an early-onset depression trajectory. Phenotypic childhood neurodevelopmental problems were markedly increased in the early-adolescent-onset group (by 5 to 7-fold) compared to the “typical” depression trajectory. Studies with follow-up further into adult life will help to clarify the adult mental health outcomes of these groups.

Strengths of this study include the repeated-measures longitudinal design where depression was assessed using exactly the same measure and informant. Typically, longitudinal studies include changes in measurement and informant, in particular, as children grow-up the informant tends to change from the parent to the young person themselves. This provides a challenge to studies seeking to examine the development of symptoms over time because changes of measurement and informant can affect results. This invariance of measurement over time is an important strength. A limitation is that like many longitudinal studies, ALSPAC suffers from non-random attrition over time (eTable 3). This is likely to result in conservative estimates of the prevalence of the elevated depression trajectory groups. We used a number of approaches to deal with missing data including FIML in trajectory modelling and inverse probability weighting (IPW) for tests of association. The
pattern of results replicated using IPW. It should be noted nonetheless that the missing data assumption made in our analyses is that there are not systematic differences between those who do and do not provide trajectory data and membership in the sample after conditioning on the other variables in the model (e.g. PRS and variables included in the IPW analysis). Depression was assessed using self-reported questionnaire rather than clinical assessment. Nonetheless, subthreshold symptoms are associated with impairment and subsequent MDD\textsuperscript{2-4}. We focused on the full sample rather than examining gender differences although results were similar when carried out separately for males and females (results available from first author). It was not possible to investigate rates of mania or bipolar disorder in the trajectory groups. However, evidence is inconsistent on the link with early-onset depression and bipolar disorder\textsuperscript{13,15,28}. The follow-up period in this study was limited to early adult life. A final limitation is that polygenic risk scores are weak predictors and only explain a small to modest proportion of phenotypic variance as they do in the present paper. However they do provide a useful biological indicator of genetic liability\textsuperscript{75}.

In summary, findings suggest etiologically distinct trajectories of depressive symptoms in youth dependent on age-at-onset and that neurodevelopmental genetic risk contributes to very early onset depression.
Figure 1. Developmental trajectories of depressive symptoms

- Probability of clinically significant depression
- Age (years)
- Low-risk (73.7%)
- Late-adolescent onset (17.3%)
- Early-adolescent onset (9.0%)
Table 1. Associations of polygenic risk scores with trajectory classes

<table>
<thead>
<tr>
<th></th>
<th>Early-adolescent-onset (9.0%)</th>
<th>95% CI</th>
<th>p</th>
<th>Late-adolescent-onset (17.3%)</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Univariate associations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDD PRS</td>
<td>1.24</td>
<td>1.06, 1.46</td>
<td>.007</td>
<td>1.27</td>
<td>1.09, 1.48</td>
<td>.003</td>
</tr>
<tr>
<td>Schizophrenia PRS</td>
<td>1.22</td>
<td>1.04, 1.43</td>
<td>.013</td>
<td>.95</td>
<td>.82, 1.11</td>
<td>.555</td>
</tr>
<tr>
<td>ADHD PRS</td>
<td>1.32</td>
<td>1.13, 1.54</td>
<td>&lt;.001</td>
<td>.94</td>
<td>.80, 1.11</td>
<td>.482</td>
</tr>
<tr>
<td><strong>Multivariate associations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDD PRS</td>
<td>1.16</td>
<td>.98, 1.36</td>
<td>.086</td>
<td>1.31</td>
<td>1.12, 1.53</td>
<td>.001</td>
</tr>
<tr>
<td>Schizophrenia PRS</td>
<td>1.19</td>
<td>1.01, 1.41</td>
<td>.040</td>
<td>.93</td>
<td>.79, 1.10</td>
<td>.391</td>
</tr>
<tr>
<td>ADHD PRS</td>
<td>1.27</td>
<td>1.08, 1.50</td>
<td>.003</td>
<td>.90</td>
<td>.76, 1.07</td>
<td>.229</td>
</tr>
</tbody>
</table>

Footnote to Table 1: Low class as the reference. OR= odds ratio for 1 standard deviation unit change, MDD = major depressive disorder, PRS = polygenic risk score. Late vs. early (multivariate analyses) MDD OR=1.13 (.88-1.46), p=.334; schizophrenia PRS OR=.78 (.60-1.01), p=.065, ADHD PRS OR=.71 (.55-.92), p=.009
### Table 2: phenotypic associations with trajectory class

<table>
<thead>
<tr>
<th>Classes</th>
<th>Early-adolescent-onset</th>
<th>Late-adolescent-onset</th>
<th>Difference between early-adolescent and later-adolescent onset classes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(9.0%) % or OR</td>
<td>(17.3%) % or OR</td>
<td></td>
</tr>
<tr>
<td>Gender (%)</td>
<td>74.3 &lt;.001</td>
<td>73.2 &lt;.001</td>
<td>$\chi^2 = .015 \ (t), \ p=.904$</td>
</tr>
<tr>
<td>Maternal education (completed A-levels) (%)</td>
<td>39.1 .012</td>
<td>34.9 .001</td>
<td>$\chi^2 = .707 \ (t), \ p=.440$</td>
</tr>
<tr>
<td>Childhood ADHD (%)</td>
<td>6.3 .008</td>
<td>0.9 .365</td>
<td>$\chi^2 = 6.837 \ (t), \ p=.009$</td>
</tr>
<tr>
<td>Pragmatic language difficulties</td>
<td>.63 (.55, .71) &lt;.001</td>
<td>.82 (.72, .94) .006</td>
<td>OR = 1.31, p=.004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\chi^2 = 11.709 \ (t), \ p=.001 \ (for \ cut-point)$</td>
</tr>
<tr>
<td>Social communication difficulties</td>
<td>1.50 (1.34, 1.68) &lt;.001</td>
<td>1.01 (.87, 1.18) 0.855</td>
<td>OR=.68, p&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\chi^2 = 18.819 \ (t), \ p=.001 \ (for \ cut-point)$</td>
</tr>
<tr>
<td>Psychotic experiences (12 yrs)</td>
<td>1.47 (1.35, 1.61) &lt;.001</td>
<td>0.89 (.64, 1.22) .455</td>
<td>OR=.60, p=.003</td>
</tr>
<tr>
<td>Psychotic experiences (17 yrs)</td>
<td>1.57 (1.36, 1.80) &lt;.001</td>
<td>1.54 (1.33, 1.79) &lt;.001</td>
<td>OR=0.99, p=.740</td>
</tr>
</tbody>
</table>

Footnote to Table 2: **Continuous** scores are standardized so odds ratios are for one standard deviation increase. Social communication score – higher scores represent more problems; pragmatic language score – lower scores represent more difficulties. Low group is the reference.
group except for tests of comparison between early-adolescent and later-adolescent onset groups where the early-adolescent onset group is the reference group. $\chi^2$ tests of difference for social communication and pragmatic language difficulties used the established clinical cut-points for identifying problems (eAppendix).


**Article Title:** Early onset depression: characterising developmental trajectories and the role of psychiatric genetic risk variants

**Authors:** Frances Rice, PhD ¹, Lucy Riglin, PhD ¹, Ajay K Thapar, PhD, MD ¹, Jon Heron, PhD ², Richard Anney, PhD ¹, Michael C O’Donovan, PhD, MD ¹, Anita Thapar, PhD, MD ¹

1 MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, UK
2 School of Social and Community Medicine, University of Bristol, UK

**Correspondence:**
F Rice
MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Hadyn Ellis Building, Maindy Road, Cardiff University, CF24 4HQ, UK
Email: ricef2@cardiff.ac.uk
Phone: 00 44 (0)29 20 688384

**Author contributions**
Conceived the paper: FR, AKT, LR, JH, AT
Drafted the paper: FR, LR, RA
Critically revised the paper FR, AKT, LR, JH, RA, AT, MO
Statistical analysis LR, AKT, RA, FR
Statistical advice JH
FR and RA had full access to all the data in the study and take responsibility for the accuracy of the statistical analysis

Revision date: 6th September 2018
Word count (text only): 2999

All authors approved the final version
Email addresses for other authors:
ThaparAK@cardiff.ac.uk
RiglinL@cardiff.ac.uk
Jon.Heron@bristol.ac.uk
AnneyR@cardiff.ac.uk
Thapar@cardiff.ac.uk
ODonovanMC@cardiff.ac.uk
Conflict of interest disclosures

None

Acknowledgements: This study was supported by a seedcorn grant from the MRC Centre for Neuropsychiatric Genetics and Genomics and the Medical Research Council MR/R004609/1.

Role of the Funder/Sponsor: The funding source 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.

Additional contributions
We acknowledge the members of the Psychiatric Genomics Consortium for the publicly available data used as the discovery samples in this article. We thank the Bipolar Disorder Working Group of the Psychiatric Genomics Consortium for providing the bipolar disorder summary statistics used in this study. We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

Funding
The UK Medical Research Council and Wellcome (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. A comprehensive list of grants funding is available on the ALSPAC website. This research was specifically funded by Wellcome Trust 08426812/Z/07/Z, Wellcome Trust and MRC 092731 which funded data collection on depression. This publication is the work of the authors; Frances Rice and Richard Anney will serve as guarantors for the contents of this paper. Genome-wide association study data were generated by Sample Logistics and Genotyping Facilities at
Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe.

LR is supported by the Wellcome Trust 204895/Z/16/Z. FR receives funding from the Medical Research Council MR/R004609/1.