THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Genome-wide Association for Major Depression Through Age at Onset Stratification

Citation for published version:
CONVERGE Consortium, CARDIoGRAM Consortium, GERAD1 Consortium, Power, RA, Tansey, KE, Buttenschøn, HN, Cohen-Woods, S, Bigdeli, T, Hall, LS, Kutalik, Z, Lee, SH, Ripke, S, Steinberg, S, Teumer, A, Viktorin, A, Wray, NR, Arolt, V, Baune, BT, Boomsma, DI, Børglum, AD, Byrne, EM, Castelao, E, Craddock, N, Craig, IW, Dannlowski, U, Deary, IJ, Degenhardt, F, Forstner, AJ, Gordon, SD, Grabe, HJ, Grove, J, Hamilton, SP, Hayward, C, Heath, AC, Hocking, LJ, Homuth, G, Hottenga, JJ, Kloiber, S, Krogh, J, Landén, M, Lang, M, Levinson, DF, Lichtenstein, P, Lucae, S, MacIntyre, DJ, Madden, P, Magnusson, PKE, Martin, NG, McIntosh, AM, Middeldorp, CM, Milaneschi, Y, Porteous, DJ & Thomson, P 2017, 'Genome-wide Association for Major Depression Through Age at Onset Stratification: Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium', *Biological Psychiatry*, vol. 81, no. 4. https://doi.org/10.1016/j.biopsych.2016.05.010

Digital Object Identifier (DOI):
10.1016/j.biopsych.2016.05.010

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Biological Psychiatry

Publisher Rights Statement:
Open Access funded by Medical Research Council
Under a Creative Commons license

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 25. Apr. 2024

Archival Report

Biological Psychiatry

Genome-wide Association for Major Depression Through Age at Onset Stratification: Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium

Robert A. Power, Katherine E. Tansey, Henriette Nørmølle Buttenschøn, Sarah Cohen-Woods, Tim Bigdeli, Lynsey S. Hall, Zoltán Kutalik, S. Hong Lee, Stephan Ripke, Stacy Steinberg, Alexander Teumer, Alexander Viktorin, Naomi R. Wray, Volker Arolt, Bernard T. Baune, Dorret I. Boomsma, Anders D. Børglum, Enda M. Byrne, Enrique Castelao, Nick Craddock, Ian W. Craig, Udo Dannlowski, Ian J. Deary, Franziska Degenhardt, Andreas J. Forstner, Scott D. Gordon, Hans J. Grabe, Jakob Grove, Steven P. Hamilton, Caroline Hayward, Andrew C. Heath, Lynne J. Hocking, Georg Homuth, Jouke J. Hottenga, Stefan Kloiber, Jesper Krogh, Mikael Landén, Maren Lang, Douglas F. Levinson, Paul Lichtenstein, Susanne Lucae, Donald J. MacIntyre, Pamela Madden, Patrik K.E. Magnusson, Nicholas G. Martin, Andrew M. McIntosh, Christel M. Middeldorp, Yuri Milaneschi, Grant W. Montgomery, Ole Mors, Bertram Müller-Myhsok, Dale R. Nyholt, Hogni Oskarsson, Michael J. Owen, Sandosh Padmanabhan, Brenda W.J.H. Penninx, Michele L. Pergadia, David J. Porteous, James B. Potash, Martin Preisig, Margarita Rivera, Jianxin Shi, Stanley I. Shyn, Engilbert Sigurdsson, Johannes H. Smit, Blair H. Smith, Hreinn Stefansson, Kari Stefansson, Jana Strohmaier, Patrick F. Sullivan, Pippa Thomson, Thorgeir E. Thorgeirsson, Sandra Van der Auwera, Myrna M. Weissman, CONVERGE Consortium, CARDIoGRAM Consortium, GERAD1 Consortium, Gerome Breen, and Cathryn M. Lewis

ABSTRACT

BACKGROUND: Major depressive disorder (MDD) is a disabling mood disorder, and despite a known heritable component, a large meta-analysis of genome-wide association studies revealed no replicable genetic risk variants. Given prior evidence of heterogeneity by age at onset in MDD, we tested whether genome-wide significant risk variants for MDD could be identified in cases subdivided by age at onset.

METHODS: Discovery case-control genome-wide association studies were performed where cases were stratified using increasing/decreasing age-at-onset cutoffs; significant single nucleotide polymorphisms were tested in nine independent replication samples, giving a total sample of 22,158 cases and 133,749 control subjects for subsetting. Polygenic score analysis was used to examine whether differences in shared genetic risk exists between earlier and adult-onset MDD with commonly comorbid disorders of schizophrenia, bipolar disorder, Alzheimer's disease, and coronary artery disease.

RESULTS: We identified one replicated genome-wide significant locus associated with adult-onset (>27 years) MDD (rs7647854, odds ratio: 1.16, 95% confidence interval: 1.11–1.21, $p = 5.2 \times 10^{-11}$). Using polygenic score analyses, we show that earlier-onset MDD is genetically more similar to schizophrenia and bipolar disorder than adult-onset MDD.

CONCLUSIONS: We demonstrate that using additional phenotype data previously collected by genetic studies to tackle phenotypic heterogeneity in MDD can successfully lead to the discovery of genetic risk factor despite reduced sample size. Furthermore, our results suggest that the genetic susceptibility to MDD differs between adult- and earlier-onset MDD, with earlier-onset cases having a greater genetic overlap with schizophrenia and bipolar disorder.

Keywords: Age at onset, GWAS, Heterogeneity, Major depressive disorder, Polygenic scoring, Stratification

http://dx.doi.org/10.1016/j.biopsych.2016.05.010

© 2016 Society of Biological Psychiatry. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Major depressive disorder (MDD) is a highly prevalent and heterogeneous disorder (1). With most individuals experiencing recurrent episodes throughout life (2), MDD is now the second leading cause of disability worldwide (3). MDD is defined by low mood and energy, inability to experience enjoyment, changes to eating and sleep patterns, feelings of guilt or worthlessness, and suicidal thoughts (4). Along with excess mortality and increased risk of suicide (5), MDD is associated with worse clinical outcomes when comorbid with health problems such as cardiovascular disease and cancer (6,7). Although the heritability is estimated at 31% to 42% (8), the causal variants remain elusive: a recent large mega-analysis with over 9000 MDD cases failed to identify any replicable associations (9), despite successes in similarly sized studies of schizophrenia (SCZ) and bipolar disorder (BPD) (10,11). This lack of biological markers may be among the causes for the well-established underfunding of research into MDD relative to its economic and health burden (12) and the reported stigmatization of sufferers (13).

However, several differences exist between MDD and other psychiatric disorders in which replicable genetic associations have been identified, including higher prevalence, greater diagnostic uncertainty, lower heritability, and, crucially, increased heterogeneity. One known source of heterogeneity that may contribute substantially is age at onset (AAO). Onset can occur at any stage of life, yet many factors associated with MDD are either age specific or age restricted. These include biological events such as puberty, menopause, and dementia, and environmental risk factors including childhood maltreatment, childbirth, and divorce. Earlier onset is associated with increased risk in first-degree relatives and with higher heritability (14–18). Considerable differences in the transmission of early- versus late-onset MDD have also been reported (19,20), with some studies suggesting the effects of novel genetic risk factors for MDD appearing later in life (21,22).

Here we build on the previous mega-analysis of the Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium (PGC-MDD), using AAO to stratify cases within a sample of 8920 cases and 9521 control subjects, with the goal of reducing heterogeneity. For each of the nine PGC-MDD samples (23–30), cases were ordered by AAO within study and divided into eight groups (octiles). Genome-wide association analysis of cases in these octiles was performed systematically against control subjects for 1,235,109 autosomal single nucleotide polymorphisms (SNPs). We examined three analytic strategies: 1) genetic variants specific to early-onset MDD; 2) risk variants specific to late-onset MDD; and 3) restricting to the intermediate four octiles excluding the 25% of cases at either extreme of AAO to test for potential heterogeneity introduced from very early or very late onset. Significantly associated SNPs were taken forward for replication in nine studies comprising 13,238 cases and 124,230 control subjects. We also examined the differences between early- and late-onset MDD in their shared heritability with commonly comorbid disorders of schizophrenia, bipolar disorder, Alzheimer's disease, and coronary artery disease to identify differences in genetic etiologies across onset groups.

METHODS AND MATERIALS

Description of Samples

Full details of the studies that form the PGC-MDD are given in the supplementary materials of the original data analysis (9). Briefly, these nine MDD studies (23–30) conducted genome-wide genotyping on individual subjects of European ancestry. Subjects were required to have diagnoses of DSM-IV lifetime MDD established using structured diagnostic instruments from direct interviews by trained interviewers or clinician-administered DSM-IV checklists. Two studies required recurrent MDD and one study required recurrent, early-onset MDD. Studies ascertained cases mostly from clinical sources, and control subjects were largely randomly selected from the population and screened for lifetime history of MDD. This led to a total of 9238 cases and 9521 control subjects with genotype information.

AAO Phenotype

AAO was defined as the age at which individuals first had symptoms that met the criteria of MDD and was self-reported in all studies. Of the original 9238 cases included in the sample, 8920 (95.6%) had a reported AAO. Cases reporting an AAO older than the recorded age at interview were removed from the analysis (n = 17). Within each study, cases were ordered by AAO and then divided into octiles, giving approximately 1000 cases per octile. Octiles were defined within each study to account for differences in case ascertainment. We noted a wide range in AAO between studies ascertaining recurrent depression using the same instrument, indicating that the precise setting (study, clinic, country) was important; we therefore chose to order cases by AAO within each study, rather than across studies, or by absolute AAO cutoffs. This strategy will identify genetic variants that were specific to early or late onset, relative to the mean AAO of the recruited cases. For secondary analysis of sex-specific effects, octiles were additionally defined within only male and female subjects for analysis of sex-specific effects, and for recurrent depression that has a higher heritability. These octiles will be referred to as O1 to O8, with O1 representing the earliest onset octile and O8 representing the latest onset octile. The GenRED (Genetics of Recurrent Early-Onset Depression) study only recruited MDD cases with an AAO below 31 years (25). We compared the distribution of AAO in GenRED with other similar studies (STAR*D [Sequenced Treatment Alternatives to Relieve Depression], RADIANT-UK) and estimated that GenRED recruited the youngest cases (62% of all possible MDD cases), with older cases absent from the study (Supplemental Figure S1). GenRED cases were ordered by AAO and assigned to the appropriate octiles O1 to O5, with no cases present in O6 to O8.

Quality Control

Genotyping was described in the supplementary materials in the original analysis (11). All samples were genotyped with SNP arrays of at least 200,000 SNPs. SNPs were removed for missingness >0.02, case-control difference in SNP missingness >0.02, SNP frequency difference from HapMap3 (phase three of the International HapMap Project) >0.15, or deviation

from Hardy–Weinberg equilibrium in control subjects (exact $p < 1.0 \times 10^{-6}$). Subjects were removed for excessive missingness (>0.02), for being identical or closely related to any subject in any sample ($\hat{p} > .2$ based on common autosomal SNPs), or if there was evidence for diverging ancestry. Ancestry was estimated using multidimensional scaling applied to 8549 SNPs directly genotyped in all samples and in approximate linkage equilibrium. Imputation was performed using Beagle 3.0.440 (31) with the CEU HapMap3 data (32) to impute 1,235,109 autosomal SNP allele dosages. The first 20 ancestry-informative principal components were included as covariates, along with an indicator for each study.

Genome-wide Association Analysis

Genome-wide association analysis was performed in PLINK using logistic regression to test the association between case-control phenotype and imputed SNP dosages under an additive model (33). Genotyping coordinates are given in NCBI Build 36/UCSC hg18 (National Center for Biotechnology Information, Bethesda, MD). Quality control was conducted separately for each sample. To test for SNPs associated with distinct aspects of MDD based on AAO, we performed three hypothesis-driven analyses (Supplemental Figure S2). The first analysis targeted those SNPs associated with early-onset MDD in a series of genome-wide association analyses, initially looking at the earliest onset cases (O1) against all control subjects, then the combined O1 and O2 cases against control subjects, then O1 to O3, etc., until all cases were included. This approach was based on the sequential additions method (34), which was developed to account for a quantitative trait that is measured in cases but not control subjects and provide an estimate of the best phenotype definition for future studies. The second analysis was similarly performed to examine those SNPs associated with later-onset MDD, but reversing the procedure (i.e., first looking at the latest-onset cases [O8] against all control subjects, then O7–O8 against control subjects, then O6–O8, etc., until all cases were included). The third analysis tested whether the extremes of AAO, both early and late, were introducing heterogeneity to the cases excluding O1 to O2 and O7 to O8, leaving those cases with onset within the interquartile range of AAO for each study (O3–O6). We then tested O3 to O6 cases against all control subjects. Each analysis was performed using all cases, male cases only, female cases only, and recurrent cases only. Analyses of all, or almost all, cases (O1–O8, O1–O7, or O2–O8) were used to identify SNPs that reached greatest significance without an AAO-specific effect and so were omitted from further analysis as would have been captured in the primary analysis of this data. In total, 52 genome-wide analyses were performed [$= (6 + 6 + 1) \times 4$], making the standard multiple testing threshold of $p < 5.0 \times 10^{-8}$ for genome-wide significance anticonservative. We applied a Bonferroni correction for 52 analyses to genome-wide significance, which is highly conservative because many analyses were highly correlated. For replication, we selected SNPs with $p < 9.5 \times 10^{-10}$ in the discovery sample in more than one analysis (either by sex or recurrence or inclusion of octiles); the combination of cases that yielded the greatest significance was chosen as the basis for replication. This analysis strategy uses AAO as a stratifying variable to construct subsets of cases that may be more homogeneous and identify SNPs that are associated with susceptibility to MDD with a restricted AAO; it does not identify SNPs that control MDD AAO, which would require a case-only analysis of AAO as a quantitative trait.

Replication Analysis

Five replication samples used in the primary analysis of this dataset had AAO information available [TwinGene (35), GenREDII/DepGenesNetwork (25), deCODE (9), PsychCoLaus (36), SHIP-LEGEND (Study of Health in Pomerania–Life-Events and Gene-Environment Interaction in Depression) (37)]. The GenPod/NEW-MEDS and Harvard i2b2 studies, which appeared in the replication of the primary analysis of this dataset, did not have AAO data available and were not included. Four new replication studies were available: 1) a collection of samples available through the University of Münster (38,39); 2) a combination of RADIANT cases from Denmark (40), the Danish DEMO and PRISME studies of MDD (41,42), and a set of Danish population control subjects; 3) the CONVERGE (China Oxford and VCU Experimental Research on Genetic Epidemiology) study of MDD cases and control subjects recruited in China (18,43,44); and 4) the Generation Scotland study, which included measures on MDD (45). These are outlined, alongside the definitions of AAO, in Supplemental Methods. Due to the early median AAO in the GenREDII/DepGenesNetwork, an artificial "median" was introduced at age 27 years, based on the median for the discovery samples. Those SNPs that passed our threshold of $p < 9.5 \times 10^{-10}$ in the discovery sample were tested for association within these nine replication studies. Due to differences in availability of genome-wide genotype data, each study was genotyped and/or imputed separately. A fixed-effect inverse variance based meta-analysis of the replication studies was performed using METAL (46).

Polygenic Analysis

We also examined the association between early- and late-onset MDD and polygenic risk scores for other psychiatric disorders, as this might reflect either shared genetic etiology or phenotypic contamination (47). Polygenic risk scores for schizophrenia (9379 cases and 7736 control subjects) and bipolar disorder (6990 cases and 4820 control subjects) were created by the PGC, using imputed data and removing overlapping control subjects, ensuring that the datasets were completely independent (10,11). Alzheimer's disease polygenic risk scores were obtained from the GERAD (Genetic and Environmental Risk for Alzheimer's disease) Consortium (48) and the coronary artery disease scores from the CARDIoGRAM Consortium (49). The imputed GERAD sample comprised 3177 cases and 7277 control subjects, and the CARDIoGRAM consortium consisted of 22,233 cases and 64,762 control subjects. All four disorders were chosen for their previous genetic and epidemiological evidence of overlap with MDD (50–52).

Polygenic risk scores were calculated for MDD cases and control subjects, summing the number of risk alleles carried, weighted by the natural log of their odds ratio in the original genome-wide association study (GWAS). Score SNPs with low minor allele frequency (<0.02) or in the major histocompatibility complex region were removed, and score datasets were

pruned for linkage disequilibrium using the clumping command in PLINK to remove SNPs within 500 kb and $r^2 > .25$ with a more significantly associated SNP. Seven scores were calculated, using a p value threshold (P_T) to restrict to the most significant SNPs in their original genome-wide association analysis ($P_T < .01, .05, .1, .2, .3, .4, .5$). Logistic regression was used to test for association between polygenic risk scores and case-control status using four different case-control sets (all MDD cases vs. control subjects, O1–O3 cases vs. control subjects, O6–O8 cases vs. control subjects, and O1–O3 cases vs. O6–O8 cases) with 20 population principal components and study indicators as covariates. We calculated the proportion of variance explained (Nagelkerke's R^2) by subtraction of a full model (covariates + polygenic risk score) from a reduced model (covariates only). The GenRED sample was not included in the analysis as they had no cases within the O6–O8 analysis; the Bonn-Mannheim study was not included in the analysis of Alzheimer's disease or coronary artery disease due to overlapping control subjects with the respective consortia.

RESULTS

Summary of AAO

After quality control, the sample consisted of 8920 MDD cases with AAO information and 9519 control subjects with 1,235,109 SNPs. The median AAO across all cases excluding GenRED was 27 years old (interquartile range: 18, 38) and the mean 28.9 ±13.64, reflecting the long tail of older onset cases (Figure 1 and Supplemental Figure S3). Median AAO by study ranged from 20 to 37 years (16 for GenRED, which recruited only recurrent cases with onset no greater than 30), with the German samples having a slightly older onset. Mean AAO was lower for recurrent than nonrecurrent cases, 27.1 to 33.0 ($p < .001$, correcting for study and excluding GenRED). AAO was also older for male subjects with a mean of 29.0 compared with 26.7 in female subjects ($p < .001$).

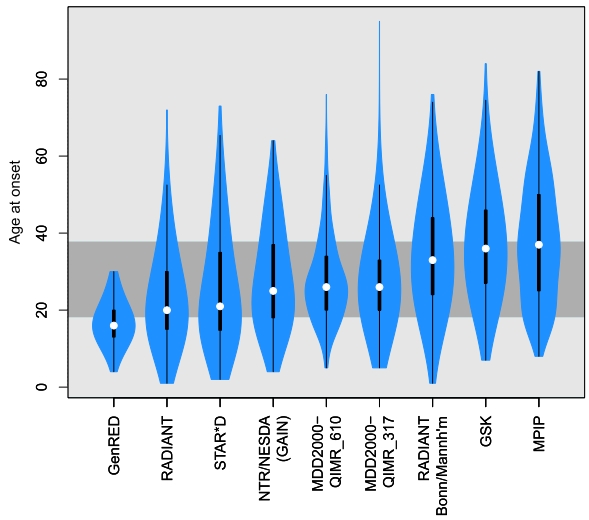

Figure 1. Distribution of age at onset across the nine studies included in the discovery analysis. Mid-gray band shows interquartile range across all studies excluding GenRED (Genetics of Recurrent Early-Onset Depression), which recruited only cases onset at 30 years or less. GAIN, Genetic Association Information Network; GSK, GlaxoSmithKline; MPIP, Max Planck Institute of Psychiatry; NESDA, Netherlands Study of Depression and Anxiety; NTR, Netherlands Twin Register.

Genome-wide Association Analysis

Our analysis of both early- and late-onset octiles and the intermediate median of AAO excluding extreme-AAO cases led to four tests passing our significance cutoff for replication ($p < 9.5 \times 10^{-10}$), all for associations with SNP rs7647854 on chromosome 3. The only genome-wide significant association for this SNP was in the 50% oldest onset cases against all control subjects (O5–O8: $p = 3.4 \times 10^{-11}$) (Supplemental Figures S4 and S5). As a secondary analysis of this SNP, we split cases within our discovery sample into nonoverlapping quartiles and analyzed them against control subjects. This showed a strong association in the oldest quartile (O7–O8: $p = 9.0 \times 10^{-10}$, odds ratio [OR]: 1.37, 95% confidence interval [CI]: 1.23–1.51) (Figure 2), a moderate effect in the third quartile (O5–O6: $p = 2.0 \times 10^{-5}$, OR: 1.23), and no evidence of association in the two youngest quartiles (O1-O2: $p = .07$, OR: 1.09; O3-O4: $p = .30$, OR: 1.06). Furthermore, when the inclusion threshold for AAO was increased 1 year at a time within each study (until only 100 cases remained), we observed a gradual increase in effect size for rs7647854 in all studies as the analysis was restricted to progressively later onset cases (Supplemental Figure S6). Analysis by sex or recurrence revealed no additional findings.

Replication Analysis

Given the results from our discovery dataset, rs7647854 was taken forward for replication in nine independent studies (Table 1). The SNP was tested for association with oldest half of MDD cases (O5–O8), because this analysis attained the smallest p value in the discovery studies. The SNP was either directly genotyped or imputed at high confidence across all studies (Supplemental Table S1). The SNP was significantly associated with MDD in a meta-analysis of these studies, with a p value of 7.5×10^{-4} and an OR of 1.10 (total number of MDD cases = 6107 and total number of control subjects 124,230) (Table 2; Supplemental Figure S5). Meta-analysis of the combined discovery sample with individual replication studies gave a p value of 5.2×10^{-11} and an OR of 1.16 (95% CI: 1.11–1.21), surpassing genome-wide significance.

Polygenic Analysis of Comorbid Illnesses

In the PGC-MDD discovery studies, polygenic risk scores for BPD and SCZ were significantly associated with early-onset MDD. Restricting to only early-onset cases (O1-O3) versus control subjects, the amount of phenotypic variability explained (BPD: $R^2 = .41\%$, $p = 1.4 \times 10^{-12}$; SCZ: $R^2 = .67\%$, $p = 3.0 \times 10^{-19}$) was much greater than for later onset (O6–O8) cases (BPD: $R^2 = .16\%$, $p = 1.9 \times 10^{-5}$; SCZ: $R^2 = .14\%$, $p = 3.9 \times 10^{-5}$) (Figure 3). A similar increase in association with late-onset comorbid disorders was not seen. Polygenic risk scores for coronary artery disease from the CARDIoGRAM Consortium (49) were weakly positively associated with MDD, but this was consistent across early and late onset cases, in contrast to BPD and SCZ (coronary artery disease: O1–O3 cases vs. control subjects, $R^2 = .05\%$, $p = .01$; O6–O8 cases vs. control

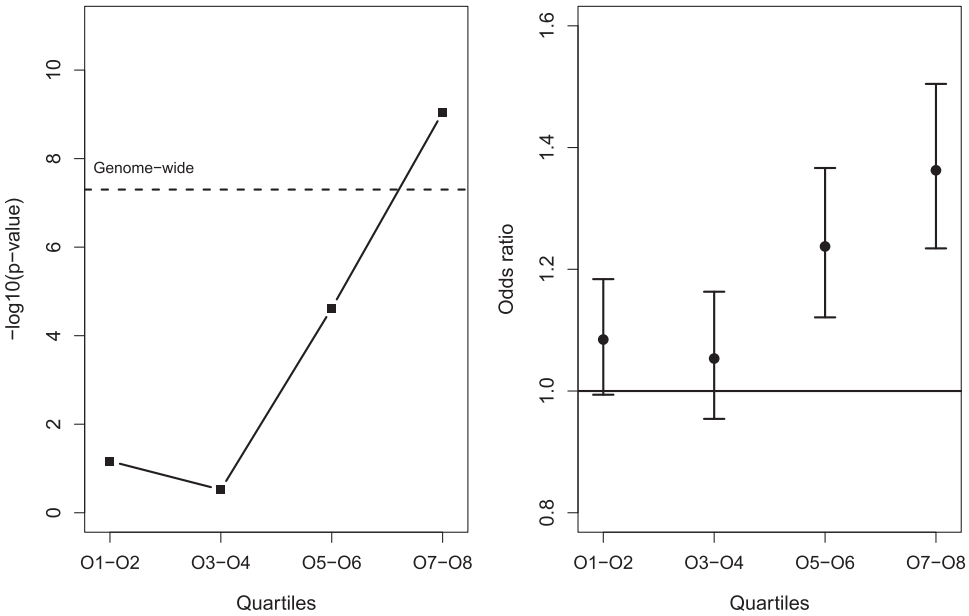

Figure 2. Evidence for association and effect size for rs7647854 on chromosome 3, with cases split into nonoverlapping quartiles by age at onset within discovery studies. O, octile.

subjects, $R^2 = .05\%$, $p = .01$; O1–O3 cases vs. O6–O8 cases, $R^2 \leq .01\%$, $p = .76$) (Figure 3). No association with MDD status was seen for the scores generated for Alzheimer's disease from the GERAD Consortium (48) regardless of AAO (O1–O3 cases vs. control subjects, $R^2 < .01\%$, $p = .868$; O6–O8 cases vs. control subjects, $R^2 = .02\%$, $p = .223$; O1–O3 cases vs. O6–O8 cases, $R^2 = .03\%$, $p = .157$) (Figure 3). Full results from these analyses are available in Supplemental Tables S2 and S3.

Table 1. Summary of MDD Discovery and Replication Cohorts

	Study	Country	Measure of AAO	Cases With AAO	Control Subjects	Median AAO
Discovery	NESDA/NTR (GAIN)	The Netherlands	CIDI	1675	1765	25
	GenRED	United States	DIGS3	1020	1253	16
	GSK	Germany	SCAN	887	864	36
	MDD2000-QIMR_610	Australia	CIDI/SSAGA	432	751	26
	MDD2000-QIMR_317	Australia	CIDI/SSAGA	1015	960	26
	MPIP	Germany	Asked at interview	373	537	37
	RADIANT Bonn/Mannheim	Germany	SCAN	883	1290	33
	RADIANT	United Kingdom	SCAN	1407	1588	20
	STAR*D	United States	Asked at interview	1228	511	21
	Total			8920	9519	
Replication	TwinGene	Sweden	SALT	1009	8601	40
	PsyCoLaus	Switzerland	DIGS	1358	1687	33
	SHIP-LEGEND	Germany	M-CIDI	381	1827	37
	GenRED2/DepGenesNetworks	United States	DIGS3	1296	930	17
	University of Münster	Germany	SCID	402	516	27
	Combined Danish sample	Denmark	SCAN	461	1197	31
	CONVERGE	China		5715	5537	34
	deCODE	Iceland	Hospital records	1005	99,175	39
	Generation Scotland	United Kingdom	SCID	1611	4760	30
	Total			13,238	124,230	
Total				22,158	133,749	

AAO, age at onset; CONVERGE, China Oxford and VCU Experimental Research on Genetic Epidemiology; CIDI, Composite International Diagnostic Interview; DIGS, Diagnostic Interview for Genetic Studies; GAIN, Genetic Association Information Network; GenRED, Genetics of Recurrent Early-Onset Depression; GSK, GlaxoSmithKline; M-CIDI, Munich-Composite International Diagnostic Interview; MDD, major depressive disorder; MPIP, Max Planck Institute of Psychiatry; NESDA, Netherlands Study of Depression and Anxiety; NTR, Netherlands Twin Register; SALT, Screening Across the Lifespan Twin; SCAN, Schedules for Clinical Assessment in Neuropsychiatry; SCID, Structured Clinical Interview for DSM Disorders; SHIP-LEGEND, Study of Health in Pomerania–Life-Events and Gene-Environment Interaction in Depression; SSAGA, Semi-Structured Assessment for the Genetics of Alcoholism; STAR*D, Sequenced Treatment Alternatives to Relieve Depression.

Table 2. Summary of Association With rs7647854, Located at 3q27.2 (186359477 Base Pairs), With Odds Ratio for the Reference Allele, G (Frequency 0.16) Compared With the Nonreference Allele A

Study	MDD Cases, *n*	Control Subjects, *n*	OR (95% CI)	*p* Value
Discovery	3869	9519	1.30 (1.20–1.40)	3.4×10^{-11}
Replication	6107	124,230	1.10 (1.05–1.17)	7.5×10^{-4}
Meta-analysis	9976	133,749	1.16 (1.11–1.21)	5.2×10^{-11}

All results reported are for the oldest half of MDD cases (O5-8), which had the strongest evidence for association in the discovery study.

CI, confidence interval; MDD, major depressive disorder; OR, odds ratio.

DISCUSSION

Our analysis of AAO and the genetic architecture of MDD suggested that AAO-specific genetic risk factors exist. Unexpectedly, the strongest associations we observed were with the oldest half of MDD cases, where previous analyses have focused on early-onset MDD, supported by studies of the genetic epidemiology of MDD. We emphasize that here "late"-onset MDD was at still a relatively young age (median onset of MDD was at age 27 years with 98% of our sample having onset before 60 years) rather than onset in old age. It is clearer to conceptualize later-onset cases here as "adult-onset" cases, and we recommend that this subgroup should be included in recruitment of future genetic studies, especially as no difference in the heritability captured by SNPs was observed for this subtype compared with early-onset cases (see Supplemental Table S3). However, substantial differences in AAO across studies make it difficult to put a precise age cutoff for this recommended adult-onset MDD.

We show significant association with rs7647854 on chromosome 3, which was associated with the 50% latest-onset cases (OR: 1.16, $p = 5.2 \times 10^{-11}$). This SNP was identified as the second strongest association in the primary mega-analysis of the PGC-MDD dataset (9), though at a much lower significance, and did not replicate (discovery $p = 6.5 \times 10^{-7}$, replication $p = .67$). Stratifying by AAO leads to much stronger associations in both the discovery and replication samples, despite the reduced sample size. Although analyzing across multiple overlapping AAO subtypes runs the risk of overfitting to maximize significance, we observed highly significant associations for this SNP in both discovery and replication samples. Furthermore, we also observed a gradual increase in effect size as both raw AAO within studies and AAO percentile across studies increased (Figure 2, Supplemental Figure S4). That this was seen both for raw AAO within studies and as a percentile, and in both European and Chinese ancestry studies across a wide variety of AAO measurement tools, suggests that whereas our cutoff of the 50% latest onset cases is arbitrary, the effect exists regardless of exact cutoff used. However, this does not rule out the possibility that AAO may be a proxy for a more homogenous subgroup of MDD based on another factor (e.g., an age-specific environmental trigger or distinct pattern of symptoms with onset later in life). That median AAO and effect size across studies was not significantly correlated ($r = .011$, $p = .68$) despite greater effect sizes by AAO percentile within studies suggests that differences in measurement of AAO across studies might obscure effects. rs7647854 is intergenic, with flanking genes including *C3orf70*, *VPS8*, *EHHADH,* and *MAP3K13. C3orf70*, *VPS8*, and *MAP3K13* all show evidence of expression in the brain in several areas of potential interest for MDD including the hypothalamus, frontal cortex, pituitary, and thyroid (GTEx; http://www.gtexportal.org/) (53). *VPS8* and *MAP3K13* also show a slight increase in RNA expression within various brain regions during neonatal development (Human Brain Transcriptome; http://hbatlas.org/) (54) with maintained expression into adulthood.

This genetic association arises with the supposedly less heritable form of adult-onset MDD, although such summary measures give no information on the effect sizes of individual SNPs. One potential explanation is greater contamination of early-onset MDD cases by individuals misclassified with MDD or having comorbid disorders. Longitudinal studies show that early-onset depressive symptoms predict not only adult depression but also psychosis [e.g., (55)], and there is significant genetic overlap between MDD and other psychiatric disorders (50). Inclusion of early-onset cases with individuals who will later develop SCZ or BPD would reduce the power of GWASs for MDD, though secondary analysis of age at interview did not support this (Supplemental Figure S7). The polygenic risk score results show that early- and adult-onset MDD cases differ in their genetic susceptibility to BPD and SCZ and suggest that some of the heterogeneity in MDD results from the inclusion of early-onset cases with a greater genetic overlap or misclassification with these disorders. We did not observe a similar pattern of association for two disorders that are often comorbid with late-onset MDD, Alzheimer's disease, and coronary artery disease (51,52). We did, however, show for the first time a genetic overlap between MDD and coronary artery disease irrespective of AAO but no overlap between MDD and Alzheimer's disease. We also found that the heritability explained by SNPs across the genome for early- and late-onset MDD did not significantly differ, suggesting, for common variants at least, that recruiting only early-onset cases would not increase power (Supplemental Table S3).

The limitations of our approach are the reduced sample size and multiple testing from stratifying cases into subtypes, and requiring potentially less reliable secondary phenotypes to be widely collected. Measures such as AAO rely on self-report and are often assessed differently across studies, which can be problematic for comparison. The analyses presented here addressed this by analyzing AAO relative to the median of a study, though this assumed that each study recruited cases from the same distribution of onset with observed differences due to how AAO was defined. Furthermore, the effect of our genome-wide associated SNP were consistent in the replication studies, ethnicities, and countries, suggesting that differences in measurement may not be as much of a limitation as expected, at least in the case of AAO in MDD. The other disadvantage of looking at more homogenous groups is the reduction in sample size. However, it has previously been shown that only modest increases in effect size may be required to offset the reduction in power from analyzing fewer cases, implying that analyses of more homogenous subgroups have the potential to identify novel associations (56).

Our study illustrates the value of using additional phenotypic information on cases in GWASs. We show here that including information on AAO increases the power to detect

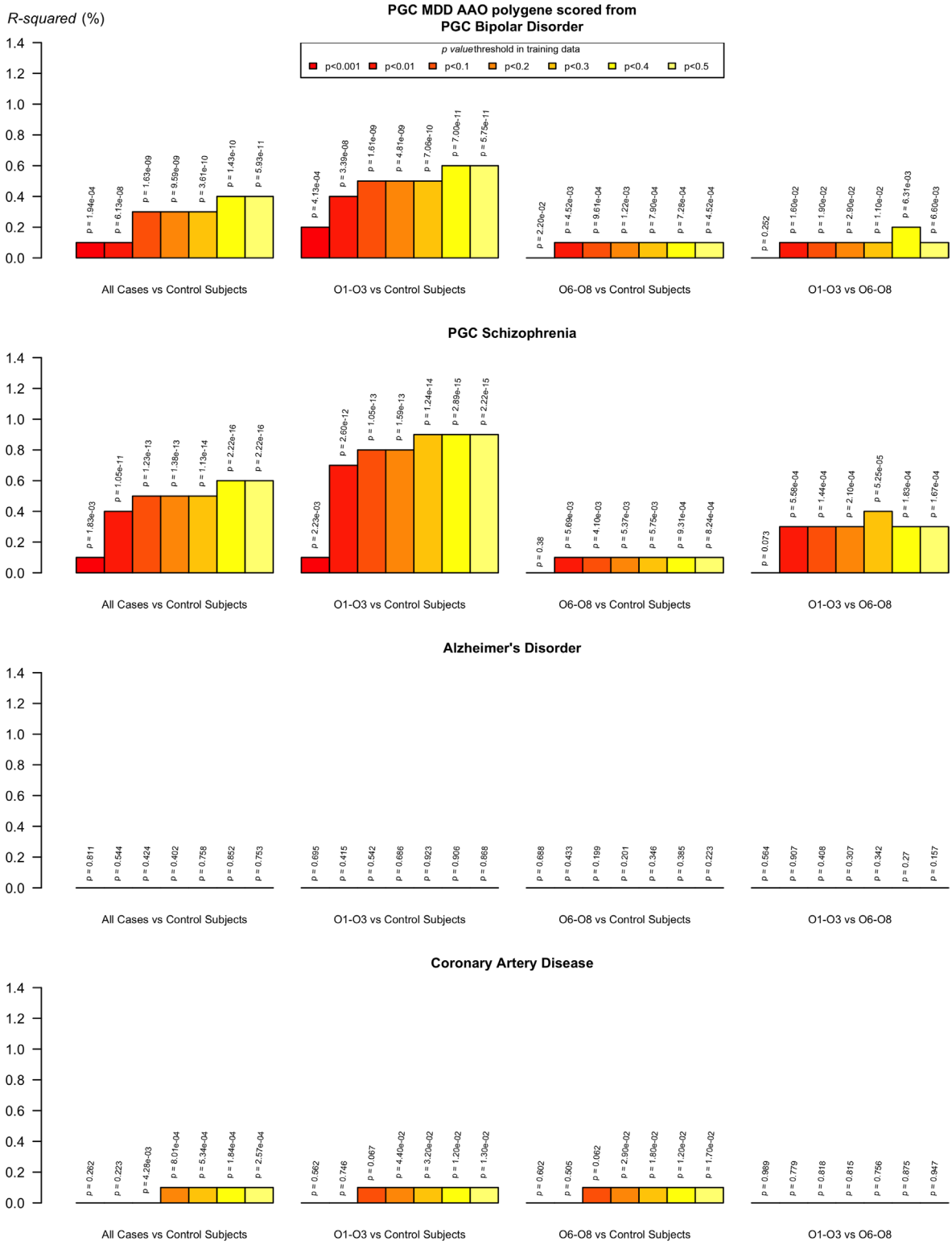

Figure 3. Polygenic risk profile scoring analysis of bipolar disorder, schizophrenia, Alzheimer's disease, and coronary artery disease within the major depressive disorder (MDD) discovery studies (excluding GenRED [Genetics of Recurrent Early-Onset Depression]). We calculated the proportion of variance explained (Nagelkerke's R^2) by subtraction of a full model (covariates + polygenic risk score) from a reduced model (covariates only). AAO, age at onset; O, octile; PGC, Psychiatric Genomics Consortium.

associations with MDD, and that analyzing polygenic risk scores from related diseases enables us to identify sources of phenotypic heterogeneity that may have hampered previous genetic studies. In contrast to other approaches that weight cases on AAO (57), our stratification of cases on AAO is agnostic to the direction of the phenotypic effect. Our approach uses the additional phenotype data previously collected by genetic studies and complements an alternative emphasis on collecting large sample sizes through minimal phenotyping. Both strategies will undoubtedly be necessary to identify and characterize different components of the genetic architecture of psychiatric disorders. Furthermore, our analysis shows that tackling phenotypic heterogeneity in MDD can successfully lead to the discovery of a genetic risk factor despite reduced sample size. The identification here of a novel genetic risk variant for MDD is of great importance due to both the scarcity of evidence for the underlying biology and its pressing economic health burden.

ACKNOWLEDGMENTS AND DISCLOSURES

The PGC was funded by National Institute of Mental Health (NIMH) Grant Nos. MH085520 (to PFS) and MH080403. Statistical analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org) hosted by SURFsara and financially supported by the Netherlands Scientific Organization Grant No. NWO 480-05-003 (to D. Posthuma) and the department of Psychology, Vrije Universiteit Amsterdam along with a supplement from the Dutch Brain Foundation. The Bonn/Mannheim GWAS was supported by the German Federal Ministry of Education and Research (BMBF) through the Integrated Genome Research Network Systematic Investigation of the Molecular Causes of Major Mood Disorders and Schizophrenia Grant Nos. 01GS08144 and 01GS08147, under the auspices of the National Genome Research Network plus, and through the Integrated Network Integrated Understanding of Causes and Mechanisms in Mental Disorders, under the auspices of the e:Med Programme Grant Nos. 01ZX1314A and 01ZX1314G. The Bonn/Mannheim GWAS was also supported by the German Research Foundation (DFG) Grant Nos. FOR2107, RI908/11-1, and NO246/10-1. The GenRED GWAS project was supported by NIMH R01 Grant Nos. MH061686 (to DFL), MH059542 (to W.H. Coryell), MH075131 (W.B. Lawson), MH059552 (JBP), MH059541 (W.A. Scheftner), and MH060912 (MMW). Max Planck Institute of Psychiatry MARS study was supported by the BMBF Program Molecular Diagnostics: Validation of Biomarkers for Diagnosis and Outcome in Major Depression by Grant No. 01ES0811. Genotyping was supported by the Bavarian Ministry of Commerce, and the BMBF in the framework of the National Genome Research Network by Grant Nos. NGFN2 and NGFN-Plus, FKZ 01GS0481 and 01GS08145. The Netherlands Study of Depression and Anxiety and the Netherlands Twin Register contributed to Genetic Association Information Network (GAIN)-MDD and to MDD2000. Funding for NTR/NESDA was from the following: the Netherlands Organization for Scientific Research (MagW/ZonMW Grant Nos. 904-61-090, 985-10- 002, 904-61-193, 480-04-004, 400-05-717, 912-100-20; Spinozapremie Grant No. 56-464-14192; Geestkracht program Grant No. 10-000-1002); the Center for Medical Systems Biology (NWO Genomics), Biobanking and Biomolecular Resources Research Infrastructure, Vrije Universiteit's Institutes for Health and Care Research and Neuroscience Campus Amsterdam, BIC/BioAssist/RK (Grant No. 2008.024); the European Science Foundation (Grant No. EU/QLRT-2001-01254); the European Community's Seventh Framework Program (Grant No. FP7/2007-2013); ENGAGE (Grant No. HEALTH-F4-2007-201413); and the European Science Council (Grant No. ERC 230374). Genotyping was funded in part by the GAIN of the Foundation for the US National Institutes of Health, and analysis was supported by grants from GAIN and the NIMH (Grant No. MH081802). Funding for the QIMR samples was provided by the Australian National Health and Medical Research Council (Grant Nos. 241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496675, 496739, 552485, 552498, 613602, 613608, 613674, 619667), the Australian Research Council (Grant Nos. FT0991360, FT0991022), the FP-5 GenomEUtwin Project (Grant No. QLG2-CT-2002-01254), and the US National Institutes of Health (Grant Nos. AA07535, AA10248, AA13320, AA13321, AA13326, AA14041, MH66206, DA12854, DA019951), and the Center for Inherited Disease Research (Baltimore, MD). RADIANT was funded by the following: a joint grant from the UK Medical Research Council and GlaxoSmithKline (Grant No. G0701420); the National Institute for Health Research Specialist Biomedical Research Centre for Mental Health at the South London and Maudsley National Health Service Foundation Trust and the Institute of Psychiatry, King's College London; the UK Medical Research Council (Grant No. G0000647), and the Marie Curie Industry-Academia Partnership and Pathways (Grant No. 286213). The GENDEP study was funded by a European Commission Framework 6 grant (EC Contract Ref.: LSHB-CT-2003-503428). Genotyping of STAR*D was supported by NIMH Grant No. MH072802 (to SPH). STAR*D was funded by NIMH Grant No. N01MH90003 to the University of Texas Southwestern Medical Center at Dallas (to A.J. Rush). The CoLaus/PsyCoLaus study was supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, and the Swiss National Science Foundation (Grant Nos. 3200B0–105993, 3200B0-118308, 33CSCO-122661, 33CS30-139468, 33CS30-148401) and two grants from GlaxoSmithKline Clinical Genetics. SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (Grant Nos. 01ZZ9603, 01ZZ0103, 01ZZ0403), the Ministry of Cultural Affairs, and the Social Ministry of the Federal State of Mecklenburg–West Pomerania. Genome-wide data have been supported by the Federal Ministry of Education and Research (Grant No. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany, and the Federal State of Mecklenburg–West Pomerania. SHIP-LEGEND is funded by the DFG (Grant No. GR 1912/5-1). The TwinGene study was supported by the Swedish Ministry for Higher Education, the Swedish Research Council (Grant No. M-2005-1112), GenomEUtwin (Grant Nos. EU/QLRT-2001-01254, QLG2-CT-2002-01254), the Swedish Foundation for Strategic Research and the US National Institutes of Health (Grant No. U01 DK066134). The collection of PRISME control subjects and genotyping of the 883 Danish control subjects was supported by grants from The Danish Strategic Research Council, The Stanley Research Foundation, and H. Lundbeck A/S. The Muenster Depression cohorts were supported by the European Union (Grant No. N Health-F2-2008-222963) and by grants from the DFG (Grant Nos. FOR 2107 and DA1151/5-1 [to UD]), Innovative Medizinische Forschung of the Medical Faculty of Münster (Grant Nos. DA120903, DA111107, and DA211012 [all to UD]). Generation Scotland is supported by a Wellcome Trust Strategic Award "Stratifying Resilience and Depression Longitudinally" (Reference No.: 104036/Z/14/Z) and core support from the Chief Scientist Office of the Scottish Government Health Directorates (Grant No. CZD/16/6) and the Scottish Funding Council (Grant No. HR03006).

The NIMH Cell Repository at Rutgers University and the NIMH Center for Collaborative Genetic Studies on Mental Disorders made essential contributions to this project. Genotyping was carried out by the Broad Institute Center for Genotyping and Analysis with support from Grant No. U54 RR020278 (which partially subsidized the genotyping of the GenRED cases). Collection and quality control analyses of the control dataset were supported by grants from NIMH and the National Alliance for Research on Schizophrenia and Depression.

We acknowledge the contributions of Dr. George S Zubenko and Dr. Wendy N Zubenko, Department of Psychiatry, University of Pittsburgh School of Medicine, to the GenRED I project. We are grateful to Knowledge Networks (Menlo Park, CA) for assistance in collecting the control dataset. We express our profound appreciation to the families who participated in this project, and to the many clinicians who facilitated the referral of participants to the study. We thank the twins and their families registered at the Australian Twin Registry for their participation in the many studies that have contributed to this research. We thank V. Mooser, G. Weaber, and P. Vollenweider who initiated the CoLaus project. We express our gratitude to the Lausanne inhabitants who volunteered to participate in the PsyCoLaus study. We would like to acknowledge the PRISME-study group, Denmark, for collection of the PRISME samples. We thank David M. Hougaard, Section of Neonatal Screening and Hormones, Statens Serum Institute, Copenhagen, Denmark; Preben Bo Mortensen, National Centre for Register-based Research, Aarhus University, Denmark; Merete Nordentoft,

Mental Health Centre, Copenhagen, Denmark; and The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Denmark. Funding from the BBSRC and MRC is gratefully acknowledged.

Data used in the preparation of this article were obtained from the Genetic and Environmental Risk for Alzheimer's disease (GERAD1) Consortium. As such, the investigators within the GERAD1 consortia contributed to the design and implementation of GERAD1 and/or provided data but did not participate in analysis or writing of this report.

SS, HS, KS, and TET are employees of deCODE Genetics/Amgen. VA received funds from the German Federal Ministry of Education and Research, from the European Union (FP 7), and from the Interdisciplinary Center for Clinical Research Münster, and he has served on the advisory boards of, or has given presentations on behalf of the following companies: Astra-Zeneca, Janssen-Organon, Lilly, Lundbeck, Servier, Pfizer, Otsuka, and Trommsdorff. BTB has received funding from the National Health and Medical Research Council Australia and honoraria from Lundbeck, Bristol-Meyers Squibb, Sanofi, Servier, Astra-Zeneca, Pfizer. IJD is supported by the MRC-BBSRC, Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (Grant No. MR/K026992/1). HJG has received funding from German Research Foundation and Federal Ministry of Education and Research Germany and speakers honoraria from Eli Lilly and Servier. CH acknowledges support from the Medical Research Council (MRC) and the Biotechnology and Biological Sciences Research Council (BBSRC). DJM is supported by an NRS Career Fellowship, funded by the Chief Scientist Office. AMM is supported by a Scottish Funding Council Senior Clinical Fellowship and by the Dame Theresa and Mortimer Sackler Foundation and has received research support from Pfizer, Janssen, and Lilly. CMM was supported by the Netherlands Organization for Scientific Research (Grant No. NOW VENI 916-76-125). BM-M has consulted for Affectis Pharmaceuticals. MP has served on the advisory boards of Lundbeck and Eli Lilly.

All other authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Institute of Psychiatry, Psychology, and Neuroscience (RAP, IWC, MR, GB, CML), King's College London, London; Medical Research Council Centre for Neuropsychiatric Genetics and Genomics (KET, NC, MJO), Institute of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, United Kingdom; Lundbeck Foundation Initiative for Integrative Psychiatric Research (HNB, DAB, JG, OM), iPSYCH; Translational Neuropsychiatry Unit (HNB), Department of Clinical Medicine, Aarhus University, Aarhus, Denmark; Discipline of Psychiatry (SC-W, BTB), School of Medicine, University of Adelaide, Adelaide, Australia; Department of Psychiatry (TB), Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, Virginia; Division of Psychiatry (LSH, DJM, AMM), University of Edinburgh, Edinburgh, United Kingdom; Institute of Social and Preventive Medicine (ZK), Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland; Queensland Brain Institute (SHL, NRW, EMB), The University of Queensland, Brisbane, Queensland; School of Environmental and Rural Science (SHL), University of New England, Armidale, New South Wales, Australia; Stanley Center for Psychiatric Research (SR), Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge; Analytic and Translational Genetics Unit (SR), Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts; Department of Psychiatry and Psychotherapy (SR), Charité, Campus Mitte, Berlin, Germany; deCODE Genetics (SS, HS, KS, TET), Reykjavik, Iceland; Institute for Community Medicine (AT), University Medicine Greifswald, Greifswald, Germany; Department of Medical Epidemiology and Biostatistics (AV, MLand, PL, PKEM, PFS), Karolinska Institutet, Stockholm, Sweden; Department of Psychiatry and Psychotherapy (VA, UD), University of Muenster, Muenster, Germany; Department of Biological Psychology (DIB, JJH, CMM), EMGO Institute for Health and Care Research and Neuroscience Campus Amsterdam, Vrije Universiteit, Amsterdam, The Netherlands; Department of Biomedicine and Centre for Integrative Sequencing (DAB, JG), iSEQ, Aarhus University, Aarhus, Denmark; Department of Psychiatry (EC, MP), Lausanne University Hospital, Lausanne, Switzerland; Department of Psychiatry (UD), University of Marburg, Marburg, Germany; Centre for Cognitive Ageing and Cognitive Epidemiology (IJD, AMM, PT); Department of Psychology (IJD), University of Edinburgh, Edinburgh, United Kingdom; Institute of Human Genetics (FD, AJF); Department of Genomics (FD, AJF), Life and Brain Center, University of Bonn, Bonn, Germany; QIMR Berghofer Medical Research Institute (SDG, NGM, GWM), Brisbane, Queensland, Australia; Department of Psychiatry and Psychotherapy (HJG, SVdA), University Medicine Greifswald, Greifswald, Germany; Department of Psychiatry (SPH), Kaiser Permanente San Francisco Medical Center, San Francisco, California; Medical Genetics Section (CH, DJP, PT), Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom; Department of Psychiatry (ACH, PM), Washington University St. Louis, St. Louis, Missouri; Division of Applied Health Sciences (LJH), University of Aberdeen, Aberdeen, United Kingdom; Interfaculty Institute for Genetics and Functional Genomics (GH), University of Greifswald, Greifswald; Max Planck Institute of Psychiatry (SK, SL, BM-M), Munich, Germany; Mental Health Center Copenhagen (JK), Mental Health Services in Capital Region, University of Copenhagen, Copenhagen, Denmark; Institute of Neuroscience and Physiology (MLand), University of Gothenburg, Gothenburg, Sweden; Department of Genetic Epidemiology in Psychiatry (MLang, JSt), Central Institute of Mental Health, Medical Faculty Mannheim/Heidelberg University, Mannheim, Germany; Department of Psychiatry and Behavioral Sciences (DFL), Stanford University, Stanford, California; Department of Psychiatry (YM, BWJHP, JHS), EMGO Institute for Health and Care Research and Neuroscience Campus Amsterdam, Vrije Universiteit Medical Center, Amsterdam, The Netherlands; Psychosis Research Unit (OM), Aarhus University Hospital, Risskov, Denmark; Munich Cluster for Systems Neurology (BM-M), Munich, Germany; Institute of Translational Medicine (BM-M), University of Liverpool, Liverpool, United Kingdom; Institute of Health and Biomedical Innovation (DRN), Queensland University of Technology, Brisbane, Queensland, Australia; Therapeia (HO), Reykjavik, Iceland; Institute of Cardiovascular and Medical Sciences (SP), University of Glasgow, Glasgow, United Kingdom; Charles E. Schmidt College of Medicine (MLP), Florida Atlantic University, Boca Raton, Florida; Department of Psychiatry (JBP), University of Iowa Carver College of Medicine, Iowa City, Iowa; Centro de Investigación Biomédica en Red de Salud Mental–Universidad de Granada e Instituto de Investigación Biosanitaria ibs (MR), Hospitales Universitarios de Granada/Universidad de Granada, Granada, Spain; Division of Cancer Epidemiology and Genetics (JSi), National Cancer Institute, National Institutes of Health, Bethesda, Maryland; Department of Psychiatry (SIS), Group Health, Seattle, Washington; Faculty of Medicine, University of Iceland (ES); Department of Psychiatry (ES), Landspitali University Hospital, Reykjavik, Iceland; Division of Population Health Sciences (BHS), University of Dundee, Dundee, United Kingdom; Department of Genetics (PFS), University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; and College of Physicians and Surgeons and the Mailman School of Public Health (MMW), Columbia University and New York State Psychiatric Institute, New York, New York.

See the Supplement for full lists of CONVERGE authors, CARDIoGRAM authors, and GERAD1 investigators.

RAP and KET contributed equally to this work.

Address correspondence to Cathryn M. Lewis, Ph.D., Institute of Psychiatry, Psychology and Neuroscience, King's College London, London SE5 8AF, United Kingdom; E-mail: cathryn.lewis@kcl.ac.uk.

Received Dec 1, 2015; revised Mar 26, 2016; accepted May 5, 2016.

Supplementary material cited in this article is available online at http://dx.doi.org/10.1016/j.biopsych.2016.05.010.

REFERENCES

1. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, *et al*. (2003): The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). JAMA-289:3095–3105.
2. Mueller TI, Leon AC, Keller MB, Solomon DA, Endicott J, Coryell W, *et al*. (1999): Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. Am J Psychiatry 156:1000–1006.

3. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, *et al*. (2012): Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010. Lancet 380:2163–2196.
4. American Psychiatric Association (1994): Diagnostic and Statistical Manual of Mental Disorders, *4th ed.* Washington, DC: American Psychiatric Association.
5. Angst F, Stassen HH, Clayton PJ, Angst J (2002): Mortality of patients with mood disorders: follow-up over 34–38 years. J Affect Disord 68: 167–181.
6. Barth J, Schumacher M, Herrmann-Lingen C (2004): Depression as a risk factor for mortality in patients with coronary heart disease: A meta-analysis. Psychosom Med 66:802–813.
7. Satin JR, Linden W, Phillips MJ (2009): Depression as a predictor of disease progression and mortality in cancer patients: A meta-analysis. Cancer 115:5349–5361.
8. Sullivan PF, Neale MC, Kendler KS (2000): Genetic epidemiology of major depression: Review and meta-analysis. Am J Psychiatry 157: 1552–1562.
9. Major Depressive Disorder Working Group of the Psychiatric Gwas Consortium, Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, *et al*. (2013): A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry 18:497–511.
10. Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium (2011): Genome-wide association study identifies five new schizophrenia loci. Nat Genet 43:969–976.
11. Psychiatric GWAS Consortium Bipolar Disorder Working Group (2011): Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. Nat Genet 43:977–983.
12. Gillum LA, Gouveia C, Dorsey ER, Pletcher M, Mathers CD, Mcculloch CE, Johnston SC (2011): NIH disease funding levels and burden of disease. PLoS One 6:e16837.
13. Sirey JA, Bruce ML, Alexopoulos GS, Perlick DA, Raue P, Friedman SJ, Meyers BS (2001): Perceived stigma as a predictor of treatment discontinuation in young and older outpatients with depression. Am J Psychiatry 158:479–481.
14. Weissman MM, Wickramaratne P, Merikangas KR, Leckman JF, Prusoff BA, Caruso KA, *et al*. (1984): Onset of major depression in early adulthood: Increased familial loading and specificity. Arch Gen Psychiatry 41:1136–1143.
15. Nierenberg AA, Trivedi MH, Fava M, Biggs MM, Shores-Wilson K, Wisniewski SR, *et al*. (2007): Family history of mood disorder and characteristics of major depressive disorder: A STAR*D (sequenced treatment alternatives to relieve depression) study. J Psychiatr Res 41:214–221.
16. Tozzi F, Prokopenko I, Perry JD, Kennedy JL, McCarthy AD, Holsboer F, *et al*. (2008): Family history of depression is associated with younger age of onset in patients with recurrent depression. Psychol Med 38:641–649.
17. Lyons MJ, Eisen SA, Goldberg J, True W, Lin N, Meyer JM, *et al*. (1998): A registry-based twin study of depression in men. Arch Gen Psychiatry 55:468–472.
18. Yang F, Li Y, Xie D, Shao C, Ren J, Wu W, *et al*. (2011): Age at onset of major depressive disorder in Han Chinese women: Relationship with clinical features and family history. J Affect Disord 135:89–94.
19. Wickramaratne PJ, Weissman MM (1998): Onset of psychopathology in offspring by developmental phase and parental depression. J Am Acad Child Adolesc Psychiatry 37, 933–942.
20. Kendler KS, Gatz M, Gardner CO, Pedersen NL (2005): Age at onset and familial risk for major depression in a Swedish national twin sample. Psychol Med 35:1573–1579.
21. Gillespie NA, Kirk KM, Evans DM, Heath AC, Hickie IB, Martin NG (2004): Do the genetic or environmental determinants of anxiety and depression change with age? A longitudinal study of Australian twins. Twin Res 7:39–53.
22. Nivard MG, Dolan CV, Kendler KS, Kan KJ, Willemsen G, Van Beijsterveldt CE, *et al*. (2015): Stability in symptoms of anxiety and depression as a function of genotype and environment: a longitudinal twin study from ages 3 to 63 years. Psychol Med 45:1039–1049.
23. Rietschel M, Mattheisen M, Frank J, Treutlein J, Degenhardt F, Breuer R, *et al*. (2010): Genome-wide association-, replication-, and neuroimaging study implicates HOMER1 in the etiology of major depression. Biol Psychiatry 68:578–585.
24. Sullivan PF, De Geus EJ, Willemsen G, James MR, Smit JH, Zandbelt T, *et al*. (2009): Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. Mol Psychiatry 14:359–375.
25. Shi J, Potash JB, Knowles JA, Weissman MM, Coryell W, Scheftner WA, *et al*. (2011): Genome-wide association study of recurrent early-onset major depressive disorder. Mol Psychiatry 16:193–201.
26. Muglia P, Tozzi F, Galwey NW, Francks C, Upmanyu R, Kong XQ, *et al*. (2010): Genome-wide association study of recurrent major depressive disorder in two European case-control cohorts. Mol Psychiatry 15:589–601.
27. Wray NR, Pergadia ML, Blackwood DH, Penninx BW, Gordon SD, Nyholt DR, *et al*. (2012): Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. Mol Psychiatry 17:36–48.
28. Shyn SI, Shi J, Kraft JB, Potash JB, Knowles JA, Weissman MM, *et al*. (2011): Novel loci for major depression identified by genome-wide association study of Sequenced Treatment Alternatives to Relieve Depression and meta-analysis of three studies. Mol Psychiatry 16, 202–215.
29. Lewis CM, Ng MY, Butler AW, Cohen-Woods S, Uher R, Pirlo K, *et al*. (2010): Genome-wide association study of major recurrent depression in the UK population. Am J Psychiatry 167:949–957.
30. Ising M, Lucae S, Binder EB, Bettecken T, Uhr M, Ripke S, *et al*. (2009): A genomewide association study points to multiple loci that predict antidepressant drug treatment outcome in depression. Arch Gen Psychiatry 66:966–975.
31. Browning BL, Browning SR (2009): A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am J Hum Genet 84:210–223.
32. International HapMap 3 Consortium, Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, *et al*. (2010): Integrating common and rare genetic variation in diverse human populations. Nature 467: 52–58.
33. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, *et al*. (2007): PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575.
34. Macgregor S, Craddock N, Holmans PA (2006):): Use of phenotypic covariates in association analysis by sequential addition of cases. Eur J Hum Genet 14:529–534.
35. Lichtenstein P, Sullivan PF, Cnattingius S, Gatz M, Johansson S, Carlstrom E, *et al*. (2006): The Swedish Twin Registry in the third millennium: An update. Twin Res Human Genet 9:875–882.
36. Preisig M, Waeber G, Vollenweider P, Bovet P, Rothen S, Vandeleur C, *et al*. (2009): The PsyCoLaus study: methodology and characteristics of the sample of a population-based survey on psychiatric disorders and their association with genetic and cardiovascular risk factors. BMC Psychiatry 9:9.
37. Volzke H, Alte D, Schmidt CO, Radke D, Lorbeer R, Friedrich N, *et al*. (2011): Cohort profile: The study of health in Pomerania. Int J Epidemiol 40:294–307.
38. Baune BT, Konrad C, Grotegerd D, Suslow T, Ohrmann P, Bauer J, *et al*. (2012): Tumor necrosis factor gene variation predicts hippocampus volume in healthy individuals. Biol Psychiatry 72:655–662.
39. Grosse L, Carvalho LA, Wijkhuijs AJ, Bellingrath S, Ruland T, Ambree O, *et al*. (2015): Clinical characteristics of inflammation-associated depression: Monocyte gene expression is age-related in major depressive disorder. Brain Behav Immun 44:48–56.
40. Uher R, Perroud N, Ng MY, Hauser J, Henigsberg N, Maier W, *et al*. (2010): Genome-wide pharmacogenetics of antidepressant response in the GENDEP project. Am J Psychiatry 167:555–564.
41. Krogh J, Petersen L, Timmermann M, Saltin B, Nordentoft M (2007): Design paper: The DEMO trial: A randomized, parallel-group, observer-blinded clinical trial of aerobic versus non-aerobic versus relaxation training for patients with light to moderate depression. Contemp Clin Trials 28:79–89.

42. Elfving B, Buttenschon HN, Foldager L, Poulsen PH, Andersen JH, Grynderup MB, *et al*. (2012): Depression, the Val66Met polymorphism, age, and gender influence the serum BDNF level. J Psychiatr Res 46: 1118–1125.
43. Yang F, Zhao H, Wang Z, Tao D, Xiao X, Niu Q, *et al*. (2014): Age at onset of recurrent major depression in Han Chinese women: A replication study. J Affect Disord 157:72–79.
44. CONVERGE Consortium (2015): Sparse whole-genome sequencing identifies two loci for major depressive disorder. Nature 523:588–591.
45. Smith BH, Campbell A, Linksted P, Fitzpatrick B, Jackson C, Kerr SM, *et al*. (2013): Cohort profile: Generation Scotland: Scottish Family Health Study (GS:SFHS): The study, its participants and their potential for genetic research on health and illness. Int J Epidemiol 42:689–700.
46. Willer CJ, Li Y, Abecasis GR (2010): METAL: Fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26:2190–2191.
47. International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, *et al*. (2009): Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 460:748–752.
48. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, *et al*. (2009): Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet 41: 1088–1093.
49. Schunkert H, Konig IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, *et al*. (2011): Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet 43:333–338.
50. Cross-Disorder Group of the Psychiatric Genomics Consortium (2013): Identification of risk loci with shared effects on five major psychiatric disorders: A genome-wide analysis. Lancet 381: 1371–1379.
51. Taylor WD, Aizenstein HJ, Alexopoulos GS (2013): The vascular depression hypothesis: Mechanisms linking vascular disease with depression. Mol Psychiatry 18:963–974.
52. Starkstein SE, Jorge R, Mizrahi R, Robinson RG (2005): The construct of minor and major depression in Alzheimer's disease. Am J Psychiatry 162:2086–2093.
53. GTEx Consortium (2013): The Genotype-Tissue Expression (GTEx) project. Nat Genet 45:580–585.
54. Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, *et al*. (2011): Spatio-temporal transcriptome of the human brain. Nature 478: 483–489.
55. Kim-Cohen J, Caspi A, Moffitt TE, Harrington H, Milne BJ, Poulton R (2003): Prior juvenile diagnoses in adults with mental disorder: Developmental follow-back of a prospective-longitudinal cohort. Arch Gen Psychiatry 60:709–717.
56. Traylor M, Markus H, Lewis CM (2015): Homogeneous case subgroups increase power in genetic association studies. Eur J Human Genet 23:863–869.
57. Zaitlen N, Lindstrom S, Pasaniuc B, Cornelis M, Genovese G, Pollack S, *et al*. (2012): Informed conditioning on clinical covariates increases power in case-control association studies. PLoS Genet 8: e1003032.