Archival Report

Biological Psychiatry

Using Clinical Characteristics to Identify Which Patients With Major Depressive Disorder Have a Higher Genetic Load for Three Psychiatric Disorders

Judith Verduijn, Yuri Milaneschi, Wouter J. Peyrot, Jouke Jan Hottenga, Abdel Abdellaoui, Eco J.C. de Geus, Johannes H. Smit, Gerome Breen, Cathryn M. Lewis, Dorret I. Boomsma, Aartjan T.F. Beekman, and Brenda W.J.H. Penninx

ABSTRACT

BACKGROUND: Limited successes of gene finding for major depressive disorder (MDD) may be partly due to phenotypic heterogeneity. We tested whether the genetic load for MDD, bipolar disorder, and schizophrenia (SCZ) is increased in phenotypically more homogenous MDD patients identified by specific clinical characteristics.

METHODS: Patients (n = 1539) with a DSM-IV MDD diagnosis and control subjects (n = 1792) were from two large cohort studies (Netherlands Study of Depression and Anxiety and Netherlands Twin Register). Genomic profile risk scores (GPRSs) for MDD, bipolar disorder, and SCZ were based on meta-analysis results of the Psychiatric Genomics Consortium. Regression analyses (adjusted for year of birth, sex, three principal components) examined the association between GPRSs with characteristics and GPRSs with MDD subgroups stratified according to the most relevant characteristics. The proportion of liability variance explained by GPRSs for each MDD subgroup was estimated.

RESULTS: GPRS-MDD explained 1.0% ($p = 4.19e^{-09}$) of MDD variance, and 1.5% ($p = 4.23e^{-09}$) for MDD endorsing nine DSM symptoms. GPRS–bipolar disorder explained 0.6% ($p = 2.97e^{-05}$) of MDD variance and 1.1% ($p = 1.30e^{-05}$) for MDD with age at onset <18 years. GPRS-SCZ explained 2.0% ($p = 6.15e^{-16}$) of MDD variance, 2.6% ($p = 2.88e^{-10}$) for MDD with higher symptom severity, and 2.3% ($p = 2.26e^{-13}$) for MDD endorsing nine DSM symptoms. An independent sample replicated the same pattern of stronger associations between cases with more DSM symptoms, as compared to overall MDD, and GPRS-SCZ.

CONCLUSIONS: MDD patients with early age at onset and higher symptom severity have an increased genetic risk for three major psychiatric disorders, suggesting that it is useful to create phenotypically more homogenous groups when searching for genes associated with MDD.

Keywords: Clinical characteristics, Heterogeneity, Genetic load, Genetics, Major depressive disorder, Replication, Staging

http://dx.doi.org/10.1016/j.biopsych.2016.05.024

Major depressive disorder (MDD) has long been recognized as heritable (~37%) (1). However, today, the largest genome-wide association study (GWAS) in MDD by the Psychiatric Genomics Consortium (PGC) has failed to find significant associations with single genetic variants (single nucleotide polymorphisms [SNPs]) (2). One likely reason why it failed is that currently available sample sizes are underpowered to detect small genetic effects (3); studies have shown that a large proportion of MDD liability is due to joint polygenic effect of common SNPs with small effects scattered across the genome and shared with other psychiatric disorders such as bipolar disorder (BIP) and schizophrenia (SCZ) (3,4). A second reason may be the clinical heterogeneity of MDD: various patients with the same diagnosis will have experienced a differential illness course with variation in, for example, experienced number, duration, and severity of episodes (5). It has therefore been suggested that GWASs should be done in phenotypically more homogenous MDD patients (2,3). The Converge Consortium showed this by examining recurrent MDD cases in Chinese women (6). However, there might be other characteristics that could be selected to enhance the genetic signal. Based on family studies (7,8), it has been suggested that the highest genetic load will be found in the most severe MDD phenotype, such as patients with young age at onset (AaO), longer (chronic) duration of symptoms, higher severity of symptoms, and recurrent episodes (1,9,10). Moreover, clinical staging strategies using jointly different clinical characteristics to define stages of MDD progression (11–13) may also be applied.

To our knowledge, it has been barely examined whether genome-wide genomic profile risk scores (GPRSs) are associated

 ISSN: 0006-3223

with clinical depression characteristics that indicate a more severe MDD phenotype. One study in depression suggests that a higher GPRS increases an individual's susceptibility for experiencing chronically high levels of depressive symptoms (14).

The current study examines whether the genetic risk for MDD, BIP, and SCZ, estimated using GPRSs generated from PGC meta-analysis results (2,15,16), is increased in phenotypically more homogenous MDD subgroups of patients stratified by clinical characteristics reflecting a more severe MDD phenotype (younger AaO, longer duration of depressive symptoms, positive MDD family history, more DSM symptoms, higher severity of depressive symptoms, and the presence of recurring MDD episodes). In addition to single characteristics, we additionally stratify patients according to an established MDD clinical staging model reflecting MDD progression (12,13). Finally, we aim to replicate the main findings in an independent dataset (17).

METHODS AND MATERIALS

Sample

The sample consisted of 3331 unrelated participants (median year of birth 1967, range 1926–1994) of North European ancestry from the NESDA (Netherlands Study of Depression and Anxiety) (n = 1851) and from the NTR (Netherlands Twin Register) (n = 1480). The methodology of both NESDA and NTR and their biobank projects have been extensively described elsewhere (18–20). The genetic sample selection is identical to the one used by Milaneschi *et al.* (21).

In short, NESDA is an ongoing longitudinal study into the onset and course of depressive and anxiety disorders. At baseline (2004–2006), 2981 adults between 18 and 65 years of age were recruited from community (19%), general practice (54%), and specialized mental health care (27%) settings to represent the entire developmental spectrum of both disorders, including healthy control subjects. After baseline assessment, 2-, 4-, and 6-year follow-up assessments were performed.

NTR has collected longitudinal data on Dutch twin families involving nearly 40,000 adult participants. The ethical review boards of contributing universities approved both studies and all participants signed informed consent.

MDD Diagnoses. The present study consisted of 1539 cases with a lifetime diagnosis of MDD (history of an MDD episode during any of their interviews) and 1792 control subjects. All cases were drawn from NESDA. The presence of MDD was assessed with the DSM-IV Composite International Diagnostic Interview (CIDI) version 2.1 (22) administered by specially trained research staff at baseline or one of the three follow-up assessments. From NESDA, we selected healthy control subjects (n = 312), who were participants without lifetime MDD or anxiety disorder.

From NTR, the majority of control subjects (n = 1480) were drawn and were participants who had no report of MDD and a low factor score based on a multivariate analyses of depressive complaints, anxiety, neuroticism, and somatic anxiety (23,24).

Clinical Characteristics. For MDD cases (all from NESDA), several clinical characteristics were assessed. AaO was ascertained via CIDI interview. Duration of depressive symptoms was examined with the Life-Chart (25) and expressed as the percentage of ~10 years (~4 years before baseline + ~6 years of follow-up) spent with depressive symptoms. Presence (yes or no) of a first-degree family member with depression was assessed with the family-tree method (26). Two different measures indexed depression severity: the highest number of DSM symptoms ever endorsed during an MDD episode extracted from the CIDI (range 5–9), and the average score on 4 measures (at each assessment) of the Inventory of Depressive Symptoms (IDS) (27). Recurring MDD episodes (yes or no) was extracted from the CIDI. Finally, we applied a clinical staging algorithm (12,13,28) (Supplement and Supplemental Figure S1), combining different clinical characteristics. Cases were assigned to one of three stages: stage 2 (n = 303) first episode; stage 3 (n = 631) recurrent/relapse episode; stage 4 (n = 605) chronic, an episode lasting longer than 2 years as indicated by the CIDI at baseline, or the Life-Chart during follow-up.

Genotyping and Genetic Relationship Matrix

Blood sample collection and DNA extraction methods have been previously described (18). Autosomal SNPs were genotyped on the Human Genome-Wide SNP Array 6.0 (Affymetrix, Santa Clara, CA) in three separate batches. Quality control (QC) steps have been previously described (29,30). Primary analyses included 497,347 SNPs. Additional stringent QC was performed to build a genetic-relationship matrix (GRM) to reduce the possibility that estimates from GRM-based analyses could be inflated by artifacts. The remaining 435,579 SNPs were used to build the GRM using GCTAv.1.24.1 (31). The QC steps are described in the Supplement.

Genomic Profile Risk Scores

As previously described (21) (more detail is available in the Supplement), results from the PGC were used to derive GPRSs for MDD (2), BIP (15), and SCZ (16). Eight sets of scored alleles were selected based on significance thresholds (Pt) (<.0001, <.001, <.005, <.01, <.05, <.1, <.5, <1) of the discovery samples associations. GPRSs were calculated as the number of scored alleles weighted by effect sizes (log-OR) from the discovery statistics (number of SNPs included for each Pt, see Supplemental Table S2). Because the GPRS construction method is based on linkage disequilibrium (LD) pruning and p thresholding, it may limit their predicting accuracy by discarding information on LD structure (32). Additionally, we derived GPRSs using the LDpred approach using LD information from a reference panel (32). Both GPRS thresholds and LDpred were standardized to a mean of 0 and standard deviation of 1 to aid interpretation of results.

Statistical Analyses

Differences in demographics between MDD cases and control subjects were examined using Mann-Whitney U test for continuous and chi-square test for categorical variables.

First, focusing on MDD cases (n = 1539), we regressed genetic risk (GPRS thresholds and LDpred) over clinical characteristics of MDD (AaO, duration of symptoms, family