Minimal phenotyping yields genome-wide association signals of low specificity for major depression

Na Cai[1,2,3*], Joana A. Revez[4], Mark J. Adams[5], Till F. M. Andlauer[6,7], Gerome Breen[8,9], Enda M. Byrne[4], Toni-Kim Clarke[5], Andreas J. Forstner[10,11,12], Hans J. Grabe[13], Steven P. Hamilton[14], Douglas F. Levinson[15], Cathryn M. Lewis[9,16], Glyn Lewis[17], Nicholas G. Martin[18], Yuri Milaneschi[19], Ole Mors[20,21], Bertram Müller-Myhsok[22,23,24], Brenda W. J. H. Penninx[19], Roy H. Perlis[25,26], Giorgio Pistis[27], James B. Potash[28], Martin Preisig[27], Jianxin Shi[29], Jordan W. Smoller[26,30,31], Fabian Streit[32], Henning Tiemeier[33,34,35], Rudolf Uher[36], Sandra Van der Auwera[13], Alexander Viktorin[37], Myrna M. Weissman[38,39], MDD Working Group of the Psychiatric Genomics Consortium[40], Kenneth S. Kendler[41], and Jonathan Flint[42]

1. Wellcome Sanger Institute, Wellcome Genome Campus, Cambridgeshire, UK
2. European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Cambridgeshire, UK
3. Helmholtz Pioneer Campus, Helmholtz Zentrum München, Neuherberg, Germany
4. Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia
5. Division of Psychiatry, University of Edinburgh, Edinburgh, UK
6. Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Munich, Germany
7. Department of Neurology, Klinikum rechts der Isar, School of Medicine, Technical University of Munich, Munich, Germany
8. NIHR Maudsley Biomedical Research Centre, King's College London, London, UK
9. Social, Genetic and Developmental Psychiatry Centre, King's College London, London, UK
10. Department of Biomedicine, University of Basel, Basel, Switzerland
11. Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Bonn, Germany
12. Centre for Human Genetics, University of Marburg, Marburg, Germany
13. Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Mecklenburg-Vorpommern, Germany
14. Department of Psychiatry, Kaiser Permanente Northern California, San Francisco, CA, USA
15. Department of Psychiatry & Behavioral Sciences, Stanford University, Stanford, CA, USA
16. Department of Medical & Molecular Genetics, King's College London, London, UK
17. Division of Psychiatry, University College London, London, UK

18. Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia
19. Department of Psychiatry, Amsterdam UMC, Vrije Universiteit and GGZinGeest, Amsterdam, The Netherlands
20. Psychosis Research Unit, Aarhus University Hospital, Risskov, Aarhus, Denmark
21. iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research, Denmark
22. Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Munich, Germany
23. Munich Cluster for Systems Neurology (SyNergy), Munich, Germany
24. University of Liverpool, Liverpool, UK
25. Department of Psychiatry, Harvard Medical School, Boston, MA, USA
26. Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA
27. Department of Psychiatry, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland
28. Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA
29. Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA
30. Psychiatric and Neurodevelopmental Genetics Unit (PNGU), Massachusetts General Hospital, Boston, MA, USA
31. Stanley Center for Psychiatric Research, Broad Institute, Cambridge, MA, USA
32. Department of Genetic Epidemiology in Psychiatry, Medical Faculty Mannheim, Central Institute of Mental Health, University of Heidelberg, Mannheim, Germany
33. Department of Epidemiology, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands
34. Department of Child and Adolescent Psychiatry, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands
35. Department of Social and Behavioral Science, Harvard TH Chan School of Public Health, Boston, MA, USA
36. Department of Psychiatry, Dalhousie University, Halifax, Nova Scotia, Canada
37. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
38. Department of Psychiatry, Columbia University, Vagelos College of Physicians and Surgeons, New York, NY, USA
39. Division of Translational Epidemiology, New York State Psychiatric Institute, New York, NY, USA

40. Individual members are listed in the Supplementary Note
41. Virginia Institute for Psychiatric and Behavioral Genetics, Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA
42. Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA, USA
*e-mail: na.cai@helmholtz-muenchen.de

Minimal phenotyping refers to the reliance on the use of a small number of self-reported items for disease case identification, increasingly used in genome-wide association studies (GWAS). Here we report differences in genetic architecture between depression defined by minimal phenotyping and strictly defined major depressive disorder (MDD): the former has a lower genotype-derived heritability that cannot be explained by inclusion of milder cases, and a higher proportion of the genome contributing to this shared genetic liability with other conditions than strictly defined MDD. GWAS based on minimal phenotyping definitions preferentially identifies loci that are not specific to MDD, and though it generates highly predictive polygenic risk scores, the predictive power can be explained entirely by large sample sizes rather than specificity for MDD. Our results show that reliance on results from minimal phenotyping may bias our views of the genetic architecture of MDD and impede our ability to identify pathways specific to MDD.

A key requisite for robust identification of genetic risk loci underlying psychiatric disease is the use of an appropriately large sample. However, the high cost of phenotyping limits sample collection[1]. One solution for reducing the burden of case identification is to use information from hospital registers[2] or individuals' self-reported symptoms, help-seeking, diagnoses or medication. We refer to the latter strategy as "minimal phenotyping", as it minimizes phenotyping costs and reduces data to a single or few self-reported answers.

However, apart from detecting more GWAS loci[3-5] (Supplementary Table 1), the consequences of sacrificing symptomatic information for genetic analyses have rarely been investigated. The consequences may be particularly important for major depressive disorder (MDD) because of its phenotypic and likely etiological heterogeneity[6], high degree of comorbidity with other psychiatric diseases[7], and substantial discrepancies between self-assessment using symptom scales and diagnoses made with full diagnostic criteria[8]. While a majority of the population self-identify as having one or two depressive symptoms at any one time, only between 9% and 20% of the population have sufficient symptoms to meet criteria for lifetime occurrence of MDD[8-10]. Furthermore, there are high rates of false positives when diagnoses are made without applying diagnostic criteria[12], and antidepressants are prescribed for a wide range of conditions other than MDD[13-15]. As such, a cohort of MDD cases obtained either through the use of self-report of illness or prescribed treatment may yield a sample that is not representative of the clinical disorder, but enriched in those with non-specific sub-clinical depressive symptoms and depression secondary to a comorbid disease.

By comparing the genetic architecture of minimal phenotyping definitions of depression with those using full diagnostic criteria for MDD in UK Biobank[16], a community-based survey of half a million men and women, we assess the implications of a minimal phenotyping strategy for GWAS in

MDD. We find that MDD defined by minimal phenotyping has a large non-specific component, and if GWAS loci from these definitions are chosen for follow-up molecular characterization, they may not be informative about biology specific to MDD.

Results

Definitions of depression in UK Biobank. We identified five ways that MDD can be defined in UK Biobank. First, self-reports of seeking medical attention for depression or related conditions provide "Help-seeking" definitions of MDD (referred to as "broad depression" in a previous GWAS[3]). Second, participants are diagnosed with "Symptom-based" MDD if, in addition to meeting help-seeking criteria, they report ever experiencing one or more of the two cardinal features of depression (low mood or anhedonia) for at least two weeks[17]. Third, a "Self-Report" definition of MDD is based on participants' self-reports of all past and current medical conditions to trained nurses. Fourth, an electronic medical record ("EMR") definition is derived from ICD10 primary and secondary illness codes in electronic health records. Finally, a "CIDI-based" diagnosis of lifetime MDD is available from subjects who answered an online "Mental Health Follow-up" questionnaire (MHQ)[18] based on the Composite International Diagnostic Interview Short Form (CIDI-SF)[19], which included DSM-5 criteria for MDD (Supplementary Note, Supplementary Fig. 1, and Supplementary Table 2). None of the definitions uses trained interviewers applying structured clinical interviews, and only the last applies operationalized criteria including symptoms, length of episode (more than two weeks) and impaired social, occupational or educational function. From hereon we refer to definitions one to three as 'minimal', the fourth as "EMR-based", and the fifth as 'strictly' defined MDD (Supplementary Note). We also included a category of participants who met the help-seeking based definition (part of "broad depression" in Howard et al.[3]) but failed to meet the symptom-based definition (as they had neither of the two cardinal symptoms of depression: depressed mood or a loss of interest or pleasure in daily activities for more than two weeks). This group we refer to as "Non-MDD" (described in detail in the Supplementary Note and Supplementary Table 3). Figure 1 outlines the different diagnostic categories and the numbers of samples that each contains.

All definitions are based on recall of episodes or symptoms of depression by participants in the UK Biobank. As priming of recall by current mood affects the reliability of such reports[20-22], we emphasize that each definition is noisy, and can be interpreted as being enriched for individuals truly fulfilling its criteria. We explore the further characteristics of all definitions and considerations in their genome-wide association analyses (GWAS) in the Supplementary Note, Supplementary Figures 2-5, and Supplementary Tables 2-11.

Minimal phenotyping definitions of depression are epidemiologically different from strictly defined MDD. We assessed whether known risk factors for MDD were similar between definitions of depression[26]. Figure 2a-g shows the mean effect (odds ratio, OR) with confidence intervals of each of the following: sex[27,28], age[29], educational attainment[30-32], socio-economic status[33], neuroticism[28,34], experience of stressful life events in the two years leading up to the baseline assessment, and cumulative traumatic life events preceding assessment[35,36] (Supplementary Note and Supplementary Table 12). Estimates of the risk factor effect sizes differed substantially, and often highly significantly, as shown by the confidence intervals in Figure 2. These may reflect differences in methods of ascertainment, or underlying pathology, between definitions of depression. Next we asked whether differences in risk factors could be used to classify definitions of depression. We applied a clustering algorithm and found that all minimal phenotyping definitions of depression cluster separately from strictly defined MDD (Fig. 2h).

Minimal definitions of depression are not just milder or noisier version of strictly defined MDD. Depression defined by minimal phenotyping has lower SNP-based heritabilities (h^2_{SNP}) than more strictly defined definitions (Fig. 3a). Self-report (SelfRepDep h^2_{SNP} = 11%, s.e. = 0.85%) and help-seeking based definitions (Psypsy h^2_{SNP} = 13%, s.e. = 1.18%; GPpsy h^2_{SNP} = 14%, s.e. = 0.81%) have heritabilities of 15% or less. By contrast, strictly defined MDD (LifetimeMSuDD) has a much higher h^2_{SNP} of 26% (s.e. = 2.15%); imposing the further criterion of recurrence brings the h^2_{SNP} up to 32% (s.e. = 2.56%). Other definitions have intermediate h^2_{SNP}. All h^2_{SNP} estimates were estimated on the liability scale using PCGCs[23] (Supplementary Note), and the trend holds regardless of the method used[23,37-39] (Supplementary Note and Supplementary Table 13). We further verified that the trend cannot be explained by potential case prevalence misestimations (Fig. 3b, Supplementary Note, Supplementary Fig. 3, and Supplementary Table 13), and was not affected by regions of high linkage-disequilibrium or complexity[40] (Supplementary Note and Supplementary Fig. 3). We compared h^2_{SNP} estimates from previous studies of MDD[4,41,42] (Supplementary Fig. 6) with our results, and found that they fit squarely into the trend we observe: the less strict the criteria used to diagnose MDD, the lower the h^2_{SNP}.

We examined the role of a number of additional factors for the lower h^2_{SNP} of minimal phenotyping definitions of MDD. First, minimal phenotyping definitions do not simply have a higher environmental contribution to MDD than the stricter definitions. When we assessed h^2_{SNP} in MDD cases with high and low exposure to environmental risk factors[44], we found that minimal phenotyping definitions of depression (GPpsy, SelfRepDep) show no significant difference between exposures, similar to or lower than strictly defined MDD (LifetimeMDD and MDDRecur) (Supplementary Note and Supplementary Table 14). Second, the minimal phenotyping definitions do not merely include milder

cases of MDD as previously hypothesized[43]. Inclusion of milder cases is equivalent to lowering the threshold for disease liability in the population above which "cases" for MDD are defined. Under the liability threshold model[45], this does not reduce the h^2_{SNP} (Supplementary Note and Extended Data Fig. 1). Instead, we show through simulations that the lower h^2_{SNP} of minimal phenotyping definitions of depression may be due to misdiagnosis of controls as cases of MDD, and misclassifications of those with other conditions as cases of MDD (Extended Data Figs. 1 and 2).

Genetic correlations between definitions of depression and other diseases. We found that the genetic correlation (rG) between minimal and strictly defined MDD includes a large proportion of non-specific liability to mental ill health. The rG between GPpsy (minimal defined MDD) and LifetimeMDD (strictly defined MDD) is 0.81 (s.e. = 0.03), significantly different than unity (Fig. 3c,d, Supplementary Table 15, Supplementary Fig. 6, and Supplementary Note). One interpretation of this finding is that the correlation represents shared genetic liability to MDD[4,5]. However, the majority of the genetic liability of LifetimeMDD due to GPpsy (approximately $rG^2 = 0.81^2 = 66\%$) is shared with the No-MDD definition, GPNoDep, as the genetic liability of GPNoDep explains approximately 70% of the genetic liability of GPpsy (rG = 0.84, s.e. = 0.05), and 34% of that of LifetimeMDD (rG = 0.58, s.e. = 0.08).

We next examined rG between different definitions of MDD and comorbid diseases, using cross-trait LDSC[46] to estimate rG with neuroticism and smoking (Extended Data Fig. 3 and Supplementary Tables 16 and 17) in UK Biobank, as well as with all psychiatric conditions in the Psychiatric Genomics Consortium (PGC)[47], including PGC1-MDD[42] and depression defined in 23andMe[4] (Supplementary Table 1). Figure 4a and Supplementary Table 18 show few differences in rG estimates between other psychiatric disorders and the different definitions of MDD in UK Biobank, consistent with previous reports[48].

Similar rG estimates can result from different genetic architectures, indexed by the extent to which genetic liability is spread across the genome. We estimated local rG_L and percentage genome contribution to total rG_T using rho-HESS[49] (Methods and Fig. 4b). 65.8% (s.e. = 0.6%), 37.1% (s.e. = 4.5%) and 42.7% (s.e. = 2.3%) of the genome explains 90% of the rG_T between strictly defined MDD (LifetimeMDD) and neuroticism, bipolar disorder, and schizophrenia, respectively. In comparison, 80.2% (s.e. = 0.6%), 47.3% (s.e. = 2.4%) and 46.8% (s.e. = 0.2%) of the genome is needed to explain the same percentage of rG_T between help-seeking based GPpsy and the same conditions (Fig. 4c). In other words, minimal phenotyping definitions of depression share more genetic loci with other psychiatric conditions than strictly defined MDD does.

Previous work[4] reported that depression defined through minimal phenotyping shows enrichment of h^2_{SNP} in regions of the genome encoding genes specifically and highly expressed in central nervous system (CNS) tissues represented in GTEx[50]. We assessed this in the definitions of depression in UK

Biobank using LDSC-SEG[51]. As shown in Figure 5, neither strictly defined MDD (LifetimeMDD) nor MDD defined based on structured clinical assessments in PGC1-MDD show significant CNS enrichments, even though larger and more heterogeneous cohorts do (Methods, Supplementary Note, Supplementary Table 1, and Extended Data Fig. 4). Notably, minimal phenotyping definition GPpsy shows a significant CNS enrichment, as does the non-MDD help-seeking definition GPNoDep, neuroticism, smoking, and other disorders in the PGC[47] such as schizophrenia[53] and bipolar disorder[54]. Our analysis shows that the degree of CNS enrichment does not relate to the strictness of the definition of MDD, and is neither sufficient nor valid evidence that any particular definition of depression better represents MDD, or captures the biological mechanisms behind MDD.

GWAS hits from minimal phenotyping are not specific to MDD. We next examined the specificity of action of individual genetic loci found in GWAS of each definition of MDD. We found that the help-seeking definitions gave the greatest number of genome-wide significant loci (27 from GPpsy and Psypsy, Supplementary Table 10) in GWAS, consistent with their larger sample sizes and statistical power for finding associations. We examined whether these loci could be detected in strictly defined MDD. Of the 27 loci from minimal phenotyping definitions, 10 showed significant effects (at $P < 0.05$ after multiple testing correction for 27 loci) on LifetimeMDD, despite the latter's much smaller sample size, consistent with the hypothesis that risk loci for minimal phenotyping MDD also act in strictly defined MDD. However, all 10 loci also showed significant effects in neuroticism, smoking, schizophrenia, or the no-MDD help-seeking condition (GPNoDep, Supplementary Table 19). Furthermore, all significant SNPs in minimal phenotyping definitions of depression have the same directions of effect on non-MDD phenotypes (Fig. 6).

We found the same pattern of results when we used loci identified from a minimal phenotyping strategy in an independent study by 23andMe that used a minimal phenotyping definition[4]. Of the 17 loci, ten replicated in GPpsy (at $P < 0.05$, after multiple testing correction for 17 loci) and three replicated in LifetimeMDD. All significant SNPs have the same directions of effect on neuroticism, smoking or schizophrenia (Extended Data Fig. 5 and Supplementary Table 20) and are therefore not specific to MDD, consistent with our analysis of minimal phenotyping definitions in UK Biobank. In summary, GWAS of minimal phenotyping definitions of depression primarily enables discovery of pathways that are shared with other conditions. It is not currently possible to assess the specificity of GWAS loci from strictly defined MDD in the same way, given the sample size of strictly defined MDD remains relatively small, and GWAS hits relatively few.

Out-of-sample prediction of MDD. Finally, we explored how well the definitions of depression in UK Biobank predict strictly defined, CIDI-based MDD in independent cohorts, using data from 23 MDD cohorts in the latest data freeze from the MDD Working Group of the Psychiatric Genomics Consortium (PGC29-MDD[5,52]; Supplementary Note, Supplementary Table 21, and Supplementary Fig. 7). We constructed polygenic risk scores (PRS) on each definition of depression in UK Biobank (Methods) and examined their prediction in each of the PGC29-MDD cohorts. Of note, PRS from all definitions of depression in UK Biobank, whether minimally or strictly phenotyped, accounted for a small proportion of variation in disease status in PGC29-MDD (Supplementary Table 22). We observed the following features.

First, PRS obtained using the full sample of GPpsy performed best at predicting MDD status in independent cohorts from PGC29-MDD (Fig. 7a, Nargelkerke's $r^2 = 0.018$, AUC = 0.56 at *P* value threshold of 0.1; Extended Data Fig. 6). However, when equal sample sizes were used (randomly down-sampled to 50,000 and case prevalence of 0.15; Methods), GPpsy no longer performed best at predicting MDD status in PGC29-MDD cohorts (Fig. 7b). Rather, PRS from the strictly defined CIDI-based MDD (LifetimeMDD) best predicted MDD disease status (Nargelkerke's $r^2 = 0.0027$, AUC = 0.52 at *P* value threshold of 0.1; Extended Data Fig. 6).

Second, the higher prediction accuracy of PRS obtained using the full sample of GPpsy can be entirely explained by its larger sample size[55] (113,260 cases, 219,362 controls, effective sample size = 298,677; Supplementary Note and Extended Data Fig. 7). We calculated the effective sample size needed for other definitions to have the same predictive power: for strictly defined LifetimeMDD, we would need an effective sample size of 129,106 (Supplementary Note and Extended Data Fig. 7), less than half of that of GPpsy.

Third, PRS from strictly defined LifetimeMDD predicted MDD disease status better in the PGC29-MDD cohorts that have a higher percentage of cases fulfilling DSM-5 symptom criteria (Supplementary Table 21 and Extended Data Fig. 8; Pearson r^2 between AUC and percentage cases in PGC29-MDD cohorts fulfilling DSM-5 symptom criteria = 0.26, $P = 0.025$, at PRS *P* value threshold = 0.1). This is consistent with the interpretation that LifetimeMDD captures signal specific to MDD. We did not observe such a trend for GPpsy (Pearson $r = 0.02$, $P = 0.57$ at PRS *P* value = 0.1) or any other definition of depression (Supplementary Table 23), suggesting their lower specificity for MDD.

Discussion

Our study demonstrates that the genetic architecture of minimal phenotyping definitions of depression is different from that of strictly defined MDD and is enriched for non-specific effects on MDD. Using a range of definitions of MDD in UK Biobank, from self-reported help seeking to a full assessment of the

DSM-5 criteria for MDD through self-reported symptoms from the MHQ, we made five key observations.

First, the heritabilities of depression defined by minimal phenotyping strategies are lower than MDD defined by full DSM-5 criteria using the CIDI questionnaire. Second, although there is substantial genetic correlation between definitions, much of the shared genetic liability is not specific to MDD, and there remain significant differences, indicating the presence of genetic effects unique to each definition. Third, a larger percentage of the genome contributes to the shared genetic liability between minimal phenotyping definitions of depression and other psychiatric conditions than those between CIDI-based MDD and other conditions, likely driven by misdiagnosis due to non-specific phenotyping. Fourth, all GWAS hits from minimal definition of depression GPpsy are shared with genetically correlated conditions such as neuroticism and smoking. Finally, while minimal phenotyping definitions enable greater predictive power for MDD status in independent cohorts, this is due to its large sample size rather than its indexing of MDD-specific effects. These results point to the non-specific nature of genetic factors identified in minimal phenotyping definitions of depression.

A number of factors need to be borne in mind when interpreting the above observations. Importantly, none of the definitions of depression in the UK Biobank were obtained from structured clinical interviews with an experienced rater (the gold standard for diagnosing MDD). The closest to that standard in UK Biobank is the online MHQ[18], based on the Composite International Diagnostic Interview Short Form (CIDI-SF)[19]. Our results suggest that self-reported diagnoses using a CIDI-SF or other diagnostic questionnaires with full DSM-5 criteria lie on the same genetic liability continuum as MDD. This would argue that MDD cases identified through self-report means using a full diagnostic questionnaire will be enriched for more strictly defined forms, with the consequence that results from genetic analysis will include loci that contribute to strictly defined MDD disease risk[64,65].

Minimal definitions of MDD do not simply include cases with lower genetic liability to MDD. This is consistent with a recent study of three large twin cohorts, which asked if a combination of MDD, depressive symptoms and neuroticism is able to capture all genetic liability of MDD[67], and showed that 65% of the genetic effects contributing to MDD are specific, and minimally defined depression (inclusive of MDD, depressive symptoms and neuroticism) can index only around one-third of the genetic liability to MDD. Similarly, previously reported high degrees of genetic correlation between MDD and depressive symptoms ($rG = 0.7$, implying roughly $rG^2 = 49\%$ of genetic factors contributing to liability of the former is attributable to that of the latter)[26] need to be put in perspective of even higher degrees of sharing between depressive symptoms and other traits such as neuroticism ($rG = 0.79$-0.94, implying roughly $rG^2 = 62$-88% of genetic variance of the former is attributable to that of the latter, especially if both were assayed at a single time point[66]).

Our findings have important implications for downstream investigations. One interpretation is that the non-specific effects found through using minimal phenotyping approaches will still advance understanding of the biology of psychiatric disorders and their treatment[5,56]. A recent report used the "quasi-replication" of GWAS loci between depressive symptoms and neuroticism as validation of their functional significance[66]. An alternative view is that these loci reflect the ways in which depressive symptoms can develop as secondary effects, including through susceptibility to adverse life events[68], personality types[28], and use or exposure to psychoactive agents like cigarette smoking[69,70]—in which case, while useful for understanding the basis of mental ill health, they are not informative about the genetic etiology of MDD, and are not useful for developing disease-specific treatment.

Our findings indicate the need for ways to integrate both strict and minimal phenotyping approaches to determine which loci to prioritize for follow-up functional analyses. They also indicate a need for means to assess symptoms for diagnosing MDD with specificity at scale, rather than reliance on minimal phenotyping. Fast and accurate diagnostic methods that use a limited number of questionnaire items are becoming available: for example, computerized adaptive diagnostic screening may be as effective for the diagnosis of MDD as an hour-long face-to-face clinician diagnostic interview[71]. There are ongoing attempts to convert behavioral health tracking data from phones or wearable devices into diagnostic information[72]. If successful, these attempts may lead to a dramatic expansion in our ability to collect data appropriate for psychiatric genetics.

Acknowledgements

We thank Omer Weissbrod, Andy Dahl, Huwenbo Shi and Verena Zuber for insightful discussions. N.C. is supported by the ESPOD Fellowship from the European Bioinformatics (EMBL-EBI) and Wellcome Sanger Institute. A.V. is supported by the Swedish Brain Foundation. C.M.L. and G.B. are funded by the National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre at South London Maudsley Foundation Trust and King's College London. In the last three years, M.M.W. has received research funds from NIMH, Templeton Foundation and the Sackler Foundation and has received royalties for publications of books on interpersonal psychotherapy from Perseus Press, Oxford University Press, on other topics from the American Psychiatric Association Press and royalties on the social adjustment scale from Multihealth Systems. The CoLaus|PsyCoLaus study was and is supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, and the Swiss National Science Foundation (grants 3200B0–105993, 3200B0-118308, 33CSCO-122661, 33CS30-139468, 33CS30-148401 and 33CS30_177535/1). The PGC has received major funding from the US National Institute of Mental Health and the US National Institute of Drug Abuse (U01 MH109528 and U01 MH1095320). This research was conducted using the UK Biobank Resource under application no. 28709, and with the

support and collaboration from all investigators who comprise the MDD Working Group of the PGC (full list in the Supplementary Note). We are greatly indebted to the hundreds of thousands of individuals who have shared their life experiences with the UK Biobank and PGC investigators.

Author Contributions

N.C. and J.F. designed the study. N.C. and J.A.R. performed the analyses. N.C. and J.F. obtained the data from the UK Biobank Resource. M.J.A., T.F.M.A., G.B., E.M.B., T.-K.C., A.J.F., H.J.G., S.P.H., D.F.L., C.M.L., G.L., N.G.M., Y.M., O.M., B.M.-M., B.W.J.H.P., R.H.P., G.P., J.B.P., M.P., J.S., J.W.S., F.S., H.T., R.U., S.V.d.A., A.V., M.M.W. and all investigators from the MDD Working Group of the PGC contributed data from the PGC. N.C., K.S.K. and J.F. interpreted the results and wrote the manuscript.

Competing interests

C.M.L. is on the scientific advisory board of Myriad Neuroscience. H.J.G. has received travel grants and speaker's honoraria from Fresenius Medical Care, Neuraxpharm and Janssen Cilag as well as research funding from Fresenius Medical Care. B.W.J.H.P. has received (non-related) research grants from Jansen Research and Boehringer Ingelheim.

References

1. Lu, J. T., Campeau, P. M. & Lee, B. H. in *Obstetrical and Gynecological Survey* (2014).
2. Ripke, S. *et al.* Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat. Genet.* **45**, 1150-1159 (2013).
3. Howard, D. M. *et al.* Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nat. Commun.* **9**, 1470 (2018).
4. Hyde, C. L. *et al.* Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat. Genet.* **48**, 1031-1036 (2016).
5. Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668-681 (2018).
6. Flint, J. & Kendler, K. S. The genetics of major depression. *Neuron* **81**, 484-503 (2014).
7. Kessler, R. C. *et al.* The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA* **289**, 3095-3105 (2003).

8. Boyd, J. H., Weissman, M. M., Thompson, W. D. & Myers, J. K. Screening for depression in a community sample. Understanding the discrepancies between depression symptom and diagnostic scales. *Arch. Gen. Psychiatry* **39**, 1195-1200 (1982).
9. Breslau, N. Depressive symptoms, major depression, and generalized anxiety: a comparison of self-reports on CES-D and results from diagnostic interviews. *Psychiatry Res.* **15**, 219-229 (1985).
10. Weissman, M. M. & Myers, J. K. Rates and risks of depressive symptoms in a United States urban community. *Acta Psychiatr. Scand.* **57**, 219-231 (1978).
11. Berardi, D. *et al.* Increased recognition of depression in primary care. Comparison between primary-care physician and ICD-10 diagnosis of depression. *Psychother. Psychosom.* **74**, 225-230 (2005).
12. Mitchell, A. J., Vaze, A. & Rao, S. Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet* **374**, 609-619 (2009).
13. Mojtabai, R. Clinician-identified depression in community settings: concordance with structured-interview diagnoses. *Psychother. Psychosom.* **82**, 161-169 (2013).
14. Druss, B. G. *et al.* Understanding mental health treatment in persons without mental diagnoses: results from the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* **64**, 1196-1203 (2007).
15. Marcus, S. C. & Olfson, M. National trends in the treatment for depression from 1998 to 2007. *Arch. Gen. Psychiatry* **67**, 1265-1273 (2010).
16. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
17. Smith, D. J. *et al.* Prevalence and characteristics of probable major depression and bipolar disorder within UK Biobank: Cross-sectional study of 172,751 participants. *PLoS ONE* **8**, e75362 (2013).
18. Davis, K. A. S. *et al.* Mental health in UK Biobank: development, implementation and results from an online questionnaire completed by 157 366 participants. *BJPsych Open* **4**, 83-90 (2018).
19. Kessler, R. C. & Ustun, T. B. The World Mental Health (WMH) Survey Initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int. J. Meth. Psych. Res.* **13**, 93-121 (2004).
20. Bromet, E. J., Dunn, L. O., Connell, M. M., Dew, M. A. & Schulberg, H. C. Long-term reliability of diagnosing lifetime major depression in a community sample. *Arch. Gen. Psychiatry* **43**, 435-440 (1986).

21. Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C. & Eaves, L. J. The lifetime history of major depression in women. Reliability of diagnosis and heritability. *Arch. Gen. Psychiatry* **50**, 863-870 (1993).
22. Rice, J. P., Rochberg, N., Endicott, J., Lavori, P. W. & Miller, C. Stability of psychiatric diagnoses. An application to the affective disorders. *Arch. Gen. Psychiatry* **49**, 824-830 (1992).
23. Weissbrod, O., Flint, J. & Rosset, S. Estimating SNP-based heritability and genetic correlation in case-control studies directly and with summary statistics. *Am. J. Hum. Genet.* **103**, 89-99 (2018).
24. Fry, A. *et al.* Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am. J. Epidemiol.* **186**, 1026-1034 (2017).
25. Adams, M. J. *et al.* Factors associated with sharing email information and mental health survey participation in large population cohorts. *bioRxiv* (2019).
26. Foley, D. L., Neale, M. C. & Kendler, K. S. Genetic and environmental risk factors for depression assessed by subject-rated symptom check list versus structured clinical interview. *Psychol. Med.* **31**, 1413-1423 (2001).
27. Kendler, K. S., Gardner, C. O., Neale, M. C. & Prescott, C. A. Genetic risk factors for major depression in men and women: Similar or different heritabilities and same or partly distinct genes? *Psychol. Med.* **31**, 605-616 (2001).
28. Kendler, K. S., Gatz, M., Gardner, C. O. & Pedersen, N. L. Personality and major depression: a Swedish longitudinal, population-based twin study. *Arch. Gen. Psychiatry* **63**, 1113-1120, (2006).
29. Alexopoulos, G. S. *et al.* 'Vascular depression' hypothesis. *Arch. Gen. Psychiatry* **54**, 915-922 (1997).
30. Kessler, R. C. *et al.* Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* **62**, 593-602 (2005).
31. Kessler, R. C., Foster, C. L., Saunders, W. B. & Stang, P. E. Social consequences of psychiatric disorders, I: Educational attainment. *Am. J. Psychiatry* **152**, 1026-1032 (1995).
32. Lorant, V. *et al.* Socioeconomic inequalities in depression: a meta-analysis. *Am. J. Epidemiol.* **157**, 98-112 (2003).
33. Kessler, R. C. Epidemiology of women and depression. *J. Affect. Disord.* **74**, 5-13 (2003).
34. Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C. & Eaves, L. J. A longitudinal twin study of personality and major depression in women. *Arch. Gen. Psychiatry* **50**, 853-862 (1993).
35. Kessler, R. C. The effects of stressful life events on depression. *Ann. Rev. Psychol.* **48**, 191-214 (1997).
36. Mazure, C. M. Life stressors as risk factors in depression. *Clinical Psychology: Science and Practice* **5**, 291-313 (1998).

37. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291-295 (2015).
38. Loh, P.-R. *et al.* Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. *Nat. Genet.* **47**, 1385-1392 (2015).
39. Shi, H., Kichaev, G. & Pasaniuc, B. Contrasting the genetic architecture of 30 complex traits from summary association data. *Am. J. Hum. Genet.* **99**, 139-153 (2016).
40. Price, A. L. *et al.* Long-range LD can confound genome scans in admixed populations. *Am. J. Hum. Genet.* **83**, 132-135 (2008).
41. CONVERGE consortium. Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* **523**, 588-591 (2015).
42. Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium *et al.* A mega-analysis of genome-wide association studies for major depressive disorder. *Mol. Psychiatry* **18**, 497-511 (2013).
43. Northern Ireland Statistics and Research Agency. 2011 Census aggregate data. UK Data Service (Edition: June 2016). (2011).
44. Peterson, R. E. *et al.* Molecular genetic analysis subdivided by adversity exposure suggests etiologic heterogeneity in major depression. *Am. J. Psychiatry* **175**, 545-554 (2018).
45. Dempster, E. R. & Lerner, I. M. Heritability of threshold characters. *Genetics* **35**, 212-236 (1950).
46. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236-1241 (2015).
47. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* **381**, 1371-1379 (2013).
48. Brainstorm Consortium *et al.* Analysis of shared heritability in common disorders of the brain. *Science* **360**, eaap8757 (2018).
49. Shi, H., Mancuso, N., Spendlove, S. & Pasaniuc, B. Local genetic correlation gives insights into the shared genetic architecture of complex traits. *Am. J. Hum. Genet.* **101**, 737-751 (2017).
50. GTEx Consortium,. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580-585 (2013).
51. Finucane, H. K. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621-629 (2018).
52. Trzaskowski, M. *et al.* Quantifying between-cohort and between-sex genetic heterogeneity in major depressive disorder. *Am. J. Med. Genet. B Neuropsychiatr. Genet*. **180**, 439-447 (2019).

53. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-427 (2014).
54. Psychiatric Genomics Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near *ODZ4*. *Nat. Genet.* **43**, 977-98, (2011).
55. Turley, P. *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat. Genet.* **50**, 229-237 (2018).
56. McIntosh, A. M., Sullivan, P. F. & Lewis, C. M. Uncovering the genetic architecture of major depression. *Neuron* **102**, 91-103 (2019).
57. Mullins, N. & Lewis, C. M. Genetics of depression: progress at last. *Curr. Psychiatry Rep.* **19**, 43 (2017).
58. Sullivan, P. F. *et al.* Psychiatric genomics: an update and an agenda. *Am. J. Psychiatry* **175**, 15-27 (2018).
59. Coyne, J. C., Schwenk, T. L. & Smolinski, M. Recognizing depression: a comparison of family physician ratings, self-report, and interview measures. *J. Am. Board Fam. Pract.* **4**, 207-215 (1991).
60. Nevin, R. L. Low validity of self-report in identifying recent mental health diagnosis among U.S. service members completing Pre-Deployment Health Assessment (PreDHA) and deployed to Afghanistan, 2007: a retrospective cohort study. *BMC Public Health* **9**, 376 (2009).
61. Clarke, D. E. *et al.* DSM-5 field trials in the United States and Canada, Part I: study design, sampling strategy, implementation, and analytic approaches. *Am. J. Psychiatry* **170**, 43-58 (2013).
62. Spitzer, R. L., Forman, J. B. & Nee, J. DSM-III field trials: I. Initial interrater diagnostic reliability. *Am. J. Psychiatry* **136**, 815-817 (1979).
63. Keller, M. B. *et al.* Results of the DSM-IV mood disorders field trial. *Am. J. Psychiatry* **152**, 843-849 (1995).
64. Corfield, E. C., Yang, Y., Martin, N. G. & Nyholt, D. R. A continuum of genetic liability for minor and major depression. *Transl. Psychiatry* **7**, e1131 (2017).
65. Direk, N. *et al.* An analysis of two genome-wide association meta-analyses identifies a new locus for broad depression phenotype. *Biol. Psychiatry* **82**, 322-329 (2017).
66. Okbay, A. *et al.* Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat. Genet.* **48**, 624-633 (2016).
67. Kendler, K. S. *et al.* Shared and specific genetic risk factors for lifetime major depression, depressive symptoms and neuroticism in three population-based twin samples. *Psychol. Med.* **49**, 2745-2753 (2018).

68. Kendler, K. S. & Karkowski-Shuman, L. Stressful life events and genetic liability to major depression: genetic control of exposure to the environment? *Psychol. Med.* **27**, 539-547 (1997).
69. Fluharty, M., Taylor, A. E., Grabski, M. & Munafo, M. R. The association of cigarette smoking with depression and anxiety: a systematic review. *Nicotine Tob. Res.* **19**, 3-13 (2017).
70. Wootton, R. E. *et al.* Causal effects of lifetime smoking on risk for depression and schizophrenia: Evidence from a Mendelian randomisation study. *bioRxiv* (2018).
71. Gibbons, R. D. *et al.* The computerized adaptive diagnostic test for major depressive disorder (CAD-MDD): a screening tool for depression. *J. Clin. Psychiatry* **74**, 669-674 (2013).
72. Freimer, N. B. & Mohr, D. C. Integrating behavioural health tracking in human genetics research. *Nat. Rev. Genet.* **20**, 129-130 (2019).

Figure legends

Figure 1 | Definitions of depression in UK Biobank. This figure shows the different definitions of MDD in UK Biobank and the color codings used consistently in this paper. For the minimal phenotyping definitions of depression presented in this paper: red for help-seeking based definitions derived from Touchscreen Questionnaire; blue for symptom-based definitions derived from Touchscreen Questionnaire; green for self-report based definition derived from Verbal Interview. For the EMR definition of depression: orange for definitions based on ICD10 codes. For strictly defined MDD: purple for CIDI-based definitions derived from Online Mental Health Followup. For the no-MDD definition: brown for GPNoDep, containing those cases in help-seeking definitions that do not have cardinal symptoms for MDD. The data fields in UK Biobank relevant for defining each phenotype are shown in "Data field in UK Biobank"; number of individuals with non-missing entries for each definition are shown in "N entries"; the qualifying answers for cases and controls respectively are shown in "Answers"; the case prevalences in each definition are shown in "Case Prevalence"; the study and definitions of depression most similar to our definitions are shown in "Most similar to". The similarities and differences between help-seeking, EMR, and symptom-based definitions with definitions of depression previously reported can be found in the Supplementary Note.

Figure 2 | Relationship between definitions of depression and environmental risk factors. a-g, These figures show forest plots of odds ratios (OR) and $-\log_{10} P$ values (LogP) between known environmental risk factors and different types (Category) of definitions of depression in UK Biobank (Definition) from logistic regression, using UK Biobank assessment center, age, sex and years of education as covariates to

control for potential geographical and demographic differences between environmental risk factors, except when they are being tested. Lifetime trauma measure was derived from Online Mental Health Followup (Supplementary Note and Supplementary Table 7); Townsend deprivation index, years of education, sex, age, recent stress and neuroticism were derived from Touchscreen Questionnaire (Supplementary Note). **h**, This figure shows a hierarchical clustering of definitions of depression in UK Biobank using ORs with environmental risk factors performed using the hclust function in R, "Height" refers to the Euclidean distance between MDD definitions at the ORs of all six risk factors. MDDRecur is not included in this clustering analysis as it is a subset of the LifetimeMDD definition. The statistics used to generate these plots are presented as Source Data.

Figure 3 | SNP-heritability and genetic correlation estimates among definitions of MDD in UK Biobank. a, This figure shows the h^2_{SNP} estimates from PCGCs[19] on each of the definitions of MDD in UK Biobank (Methods). h^2_{SNP} "h2(liab)" as shown on the figure has been converted to liability scale[44,73] using the observed prevalence of each definition of depression in UK Biobank as both population and sample prevalences (Supplementary Table 4). Error bars show the standard errors of the estimates. **b**, This figure shows the h^2_{SNP} estimates of definitions of MDD in UK Biobank from LDSC using logistic regression summary statistics on all SNPs > 5% MAF (Methods), transformed to the liability scale assuming a range of population case prevalence, from 0 to 0.5. We do not show results for case prevalence from 0.5 to 1, as they will be mirroring those from 0 to 0.5. In the figure, we indicate with a black vertical dotted line the population prevalence of 0.15, used in PGC1-MDD, and a colored vertical dotted line for the population prevalence of each definition of depression in UK Biobank. We also indicate with a black horizontal dotted line the arbitrary liability scale h^2_{SNP} of 0.2, previously estimated for MDD in PGC1-MDD. Using this, we show that at no prevalence would minimal phenotyping defined depression like GPpsy (Help-seeking definition) reach this estimate. **c**, This figure shows the genetic correlation "rG" between CIDI-based LifetimeMDD and all other definitions of MDD in UK Biobank, estimated using PCGCs. Error bars show the standard errors of the estimates. **d**, This figure shows pairwise rG between all definitions of depression in UK Biobank, also detailed in Supplementary Table 15.

Figure 4 | Genetic correlation between definitions of MDD and other psychiatric conditions. a, This figure shows the genetic correlation "rG" estimated by cross-trait LDSC[46] on the liability scale between definitions of MDD in UK Biobank with other psychiatric conditions in both UK Biobank (smoking and neuroticism) and PGC[47] (Supplementary Table 1), including schizophrenia[53] (SCZ) and bipolar disorder[54] (BIP) (Supplementary Table 1). Error bars show the standard errors of the estimates. **b**, This figure shows

the cumulative fraction of regional genetic correlation "rG" (out of sum of regional genetic correlation across all loci) between definitions of MDD in UK Biobank with SCZ in 1,703 independent loci in the genome[79] estimated using rho-HESS[49], plotted against percentage of independent loci. CIDI-based LifetimeMDD is shown in purple while help-seeking based GPpsy is shown in red. The steeper the curve, the smaller the number of loci explaining the total genetic correlation. The dotted colored curves around each solid line represent the standard errors of the estimate computed using a jackknife approach as described in Shi et al.[39]. The dotted black line represents 100% of the sum of genetic correlation between each definition of MDD in UK Biobank with SCZ. The cumulative sums of positive regional genetic correlations (right of *y* axis) go beyond 100% – this is mirrored by the negative regional genetic correlation (left of *y* axis) that go below 0%. **c**, We rank all 1,703 loci by their magnitude of genetic correlation, and ask what fraction of loci sums up to 90% of total genetic correlation. This figure shows the percentage of loci summing up to 90% of total genetic correlation "rG" between either LifetimeMDD (in purple) or GPpsy (in red) with all psychiatric conditions tested, with standard errors estimated using the same jackknife approach. The higher the percentage, the higher the number of genetic loci contributing to 90% of total genetic correlation. Error bars show the standard errors of the estimates.

Figure 5 | Tissue-specific gene expression enrichment in definitions of MDD. This figure shows the $-\log_{10}(P)$ of enrichment in h^2_{SNP} in genes specifically expressed in 44 GTEx tissues, estimated using partitioned h^2_{SNP} in LDSC; help-seeking based definitions of MDD GPpsy, as well as its constituent no-MDD phenotype GPNoDep, show enrichment of h^2_{SNP} in genes specifically expressed in CNS tissues, similar to an independent cohort of help-seeking based MDD (23andMe[4]) and other psychiatric conditions such as bipolar disorder (BIP)[54], schizophrenia (SCZ)[53], autism (AUT), personality dimension neuroticism, and behavioural trait smoking. We indicate the sample size (N) for each definition of depression and psychiatric condition.

Figure 6 | GWAS hits from minimal phenotyping definition of MDD in UK Biobank are not specific to MDD. This figure shows the odds ratios (ORs) for the risk alleles at 27 loci significantly associated with help-seeking based definitions of MDD in UK Biobank (GPpsy and Psypsy), in logistic regression GWAS conducted on CIDI (LifetimeMDD, in purple), help-seeking (GPpsy in red) and no-MDD (GPNoDep, in brown) based definitions of MDD. For comparison, we show the same in conditions other than MDD: neuroticism, smoking and SCZ (all in pink). SNPs missing in each panel are not tested in the respective GWAS. For clarity of display, scales on different panels vary to accommodate the different magnitudes of ORs of SNPs in different conditions. ORs at all 27 loci are highly consistent across phenotypes, being completely aligned in direction of effect, regardless of whether it is a definition or

MDD or a risk factor or condition other than MDD. All results are shown in Supplementary Table 14. Error bars show the standard errors of the estimates.

Figure 7 | Out-of-sample prediction of MDD in PGC cohorts. a, This figure shows the area under the curve (AUC) of polygenic risk scores (PRS) calculated for each definition of depression in UK Biobank and MDD status indicated in 19 PGC29-MDD cohorts[5], while controlling for cohort-specific effects. PRS were calculated using effect sizes at independent (LD $r^2 < 0.1$) SNPs passing *P* value thresholds 10^{-4}, 0.001, 0.01, 0.05, 0.01, 0.2, 0.5 and 1, respectively, in GWAS performed on all definitions of depression in UK Biobank. **b**, This figure shows the same analysis performed on down-sampled data (7,500 cases, 42,500 controls) for each definition of depression.

Online Methods

Genome-wide associations. To obtain and access the difference between odds ratios of associations in different definitions of depression in UK Biobank, as well as smoking (data field 20160) and neuroticism (data field 20127), we performed logistic regression (or linear regression with --standard-beta for neuroticism) on all 5,276,842 common SNPs (MAF > 5% in all 337,198 White-British, unrelated samples) in PLINK[76] (version 1.9) with 20 PCs and genotyping array as covariates.

Estimation of SNP-heritability and genetic correlation among definitions of MDD. All estimates of h^2_{SNP} are computed with the phenotype-correlation-genotype-correlation (PCGC)[77] approach implemented with PCGCs[23], using 5,276,842 common SNPs (MAF > 5% in all 337,198 White-British, unrelated samples). LD scores at SNPs were computed with LDSC[37] in 10,000 random samples drawn from the White-British samples in UK Biobank as LD reference, and MAF at all 5,276,842 common SNPs in all 337,198 White-British samples as MAF reference. Covariates were genotyping array and 20 PCs computed using samples in each definition of MDD with flashPCA[74]. Where we stratified each definition of MDD in UK Biobank into two strata by risk factors such as sex (Supplementary Note), we computed specific PCs for each definition and strata (see also Supplementary Note and Supplementary Table 13).

Estimation of genetic correlation between definitions of MDD and other conditions. Summary statistics for other psychiatric conditions from previous GWAS studies were obtained as described in Supplementary Table 1. Association summary statistics for smoking and neuroticism in UK Biobank were generated by GWAS (Supplementary Table 15-16 and Extended Data Fig. 3). We estimated the genetic correlation between definitions of MDD in UK Biobank with each of these conditions with LDSC[46], with a LD reference panel generated with EUR individuals from 1000 Genomes[78]. To obtain regional rG, we partitioned the genome into 1,703 independent loci[79] and estimated regional rG with rho-HESS[49], using a LD reference panel generated with EUR individuals from 1000 Genomes[78]. We estimated standard errors for each regional rG and the total rG across the genome using a jackknife approach implemented in HESS[39]. To assess percentage of genome contributing to total rG, we ranked all independent loci by their absolute value of regional rG, and asked how many loci would contribute 90% of the total rG.

Enrichment of SNP-heritability in genes specifically expressed in tissues. We estimated the enrichment of h^2_{SNP} in genes specifically expressed in 44 tissues in the Genotype–Tissue Expression (GTEx)[50] project using the partitioned h^2_{SNP} framework in LDSC-SEG[49], and a LD reference panel generated with EUR individuals from 1000 Genomes[78]. We obtained tissue specific gene expression annotations in GTEx tissues from LDSC-SEG, then estimated the enrichment of h^2_{SNP} in annotations that corresponded to each of the tissues together with 52 annotations in the baseline model[80]. We report the *P* value of the one-sided test of enrichment of h^2_{SNP} in genes specifically expressed in each tissue against the baseline.

Out of sample predictions of MDD. We carried out out-of-sample prediction using individual level genotype and phenotype data from the PGC29 MDD cohorts[5]. We obtained permissions from 20 cohorts with sample sizes greater than 500, among which 17 recorded endorsement of DSM-5 criteria A for MDD (Supplementary Note and Supplementary Table 21). We obtained PRS from GWAS for each definition of depression in UK Biobank, using LD-clumped (LD $r^2 < 0.1$) independent SNPs with *P* values of associations below 8 thresholds ($P < 10^{-4}$, 0.001, 0.01, 0.05, 0.1, 0.2, 0.5 and 1), and predicted MDD status in the 20 PGC cohorts using the Ricopili pipeline[82]. We obtained Nagelkerke's r^2 between the PRS and MDD status, AUC of the prediction, and variance of MDD status explained by the PRS for each cohort. We also obtained the same measures for MDD status pulling data from all cohorts, controlling for cohort differences by including it as a covariate.

Ethical approval. This research was conducted under the ethical approval from the UKBiobank Resource under application no. 28709.

Reporting Summary. Additional information on the study design is provided in the **Life Sciences Reporting Summary**.

Data availability

Genotype and phenotype data used in this study are from the full release (imputation version 2) of the UK Biobank Resource obtained under application no. 28709. We used publicly available summary statistics from other studies downloadable from the website of Psychiatric Genomics Consortium (https://www.med.unc.edu/pgc/results-and-downloads), the references for which can be found in Supplementary Table 1. We also referenced the 2011 Census aggregate data from the UK Data Service (http://dx.doi.org/10.5257/census/aggregate-2011-2).

Methods only references

73 Bycroft, C. *et al.* Genome-wide genetic data on ~500,000 UK Biobank participants. *bioRxiv* (2017).

74 Abraham, G. & Inouye, M. Fast principal component analysis of large-scale genome-wide data. *PLoS ONE* **9**, e93766 (2014).

75 McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279-1283 (2016).

76 Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7 (2015).

77 Golan, D., Lander, E. S. & Rosset, S. Measuring missing heritability: inferring the contribution of common variants. *Proc. Natl. Acad. Sci. USA* **111**, E5272-E5281 (2014).

78 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74 (2015).

79 Berisa, T. & Pickrell, J. K. Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics* **32**, 283-285 (2016).

80 Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228-1235 (2015).

81 Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-2191 (2010).

82 Lam, M. *et al.* RICOPILI: Rapid Imputation for COnsortias PIpeLIne. *bioRxiv* (2019).