Check for updates

ARTICLE   OPEN

Body mass index interacts with a genetic-risk score for depression increasing the risk of the disease in high-susceptibility individuals

Augusto Anguita-Ruiz [1,2,3,4], Juan Antonio Zarza-Rebollo [1,5 ✉], Ana M Pérez-Gutiérrez [1,5], Esther Molina[3,5,6],
Blanca Gutiérrez [3,5,7], Juan Ángel Bellón[8,9], Patricia Moreno-Peral [8], Sonia Conejo-Cerón[8], Jose María Aiarzagüena[10],
M Isabel Ballesta-Rodríguez[11], Anna Fernández[12,13], Carmen Fernández-Alonso[14], Carlos Martín-Pérez[15], Carmen Montón-Franco[16,17],
Antonina Rodríguez-Bayón[18], Álvaro Torres-Martos [1], Elena López-Isac [1,5], Jorge Cervilla [3,5,7] and Margarita Rivera[1,3,5]

© The Author(s) 2022

Depression is strongly associated with obesity among other chronic physical diseases. The latest mega- and meta-analysis of genome-wide association studies have identified multiple risk loci robustly associated with depression. In this study, we aimed to investigate whether a genetic-risk score (GRS) combining multiple depression risk single nucleotide polymorphisms (SNPs) might have utility in the prediction of this disorder in individuals with obesity. A total of 30 depression-associated SNPs were included in a GRS to predict the risk of depression in a large case-control sample from the Spanish PredictD-CCRT study, a national multicentre, randomized controlled trial, which included 104 cases of depression and 1546 controls. An unweighted GRS was calculated as a summation of the number of risk alleles for depression and incorporated into several logistic regression models with depression status as the main outcome. Constructed models were trained and evaluated in the whole recruited sample. Non-genetic-risk factors were combined with the GRS in several ways across the five predictive models in order to improve predictive ability. An enrichment functional analysis was finally conducted with the aim of providing a general understanding of the biological pathways mapped by analyzed SNPs. We found that an unweighted GRS based on 30 risk loci was significantly associated with a higher risk of depression. Although the GRS itself explained a small amount of variance of depression, we found a significant improvement in the prediction of depression after including some non-genetic-risk factors into the models. The highest predictive ability for depression was achieved when the model included an interaction term between the GRS and the body mass index (BMI), apart from the inclusion of classical demographic information as marginal terms (AUC = 0.71, 95% CI = [0.65, 0.76]). Functional analyses on the 30 SNPs composing the GRS revealed an over-representation of the mapped genes in signaling pathways involved in processes such as extracellular remodeling, proinflammatory regulatory mechanisms, and circadian rhythm alterations. Although the GRS on its own explained a small amount of variance of depression, a significant novel feature of this study is that including non-genetic-risk factors such as BMI together with a GRS came close to the conventional threshold for clinical utility used in ROC analysis and improves the prediction of depression. In this study, the highest predictive ability was achieved by the model combining the GRS and the BMI under an interaction term. Particularly, BMI was identified as a trigger-like risk factor for depression acting in a concerted way with the GRS component. This is an interesting finding since it suggests the existence of a risk overlap between both diseases, and the need for individual depression genetics-risk evaluation in subjects with obesity. This research has therefore potential clinical implications and set the basis for future research directions in exploring the link between depression and obesity-associated disorders. While it is likely that future genome-wide studies with large samples will detect novel genetic variants associated with depression, it seems clear that a combination of genetics and non-genetic information (such is the case of obesity status and other depression comorbidities) will still be needed for the optimization prediction of depression in high-susceptibility individuals.

Translational Psychiatry (2022)12:30 ; https://doi.org/10.1038/s41398-022-01783-7

[1]Department of Biochemistry and Molecular Biology II, Faculty of Pharmacy, University of Granada, Granada, Spain. [2]Institute of Nutrition and Food Technology "José Mataix", Biomedical Research Center (CIBM), University of Granada, Granada, Spain. [3]Instituto de Investigación Biosanitaria ibs.GRANADA, Granada, Spain. [4]CIBEROBN (Physiopathology of Obesity and Nutrition CB12/03/30038), Institute of Health Carlos III (ISCIII), Madrid, Spain. [5]Institute of Neurosciences 'Federico Olóriz', Biomedical Research Center (CIBM), University of Granada, Granada, Spain. [6]Department of Nursing, Faculty of Health Sciences, University of Granada, Granada, Spain. [7]Department of Psychiatry, Faculty of Medicine, University of Granada, Granada, Spain. [8]Primary Care District of Málaga-Guadalhorce, Biomedical Research Institute of Málaga (IBIMA), Primary Care Prevention and Health Promotion Network (redIAPP), Málaga, Spain. [9]Department of Public Health and Psychiatry, Faculty of Medicine, University of Málaga, Málaga, Spain. [10]San Ignacio Health Centre, Primary Care Research Unit, Osakidetza Bilbao, Spain. [11]Federico del Castillo Health Centre, Jaén, Spain. [12]Parc Sanitari Sant Joan de Déu, Fundació Sant Joan de Déu, Barcelona, Spain. [13]CIBERESP, Centro de Investigacion Biomedica en Red de Epidemiologia y Salud Publica, Madrid, Spain. [14]Service Assistance Programs, Regional Health Management Valladolid, Valladolid, Spain. [15]Marquesado Health Centre, Servicio Andaluz de Salud, Granada, Spain. [16]Casablanca Health Centre, Aragonese Institute of Health Sciences, IIS Aragón, Zaragoza, Spain. [17]Department of Medicine and Psychiatry, University of Zaragoza, Zaragoza, Spain. [18]San José Health Centre Linares, Jaén, Spain. ✉email: jazarza@ugr.es

BACKGROUND

Depression and obesity are common conditions that tend to co-exist. Depression is the most common psychiatric disorder with more than 300 million people suffering from it. At the same time, the prevalence of obesity is a serious worldwide issue and one of the major health challenges of the 21st century [1]. The co-occurrence of both conditions has been designated as one of the most important contributors to the worldwide disability burden, further leading to major personal and public health implications as well as generating an enormous economic and social cost [2, 3]. Given the high prevalence of both disorders and their consequences, understanding the nature of their relationship is a pressing clinical problem.

There is evidence that people with depression are more likely to be obese compared to psychiatrically-healthy controls [4]. Conversely, people with obesity are also more prone to develop depression than normal-weight subjects so that the association between both conditions is bidirectional [5]. Several longitudinal meta-analyses have robustly evidenced this phenomenon, showing how obesity longitudinally increases the risk of developing depression, and vice versa [6–9]. The mechanisms underlying this association remain unclear nonetheless [6]. Behavioral, socio-cultural, psychological, and biological factors have been proposed as plausible explanation [3]. Regarding behavioral and socio-cultural factors, obesity would lead to depression due to stigma, interpersonal distress, and changes in body image; while depression would lead to obesity as a result of physical inactivity, alcohol abuse, and emotional eating [1, 6, 10, 11]. Interestingly, several biological dysregulations have been further described to derive from such behavioral alterations in both depression and obesity [3]. On the other hand, it might happen that depression and obesity share some molecular disturbances, strongly connected by alterations in the systems involved in homeostatic adjustments and the brain circuitries that integrate mood regulatory responses (e.g., the hypothalamic–pituitary–adrenal (HPA) axis, immuno-inflammatory activation, neuroendocrine regulators of energy metabolism, or the microbiome) [5, 6, 12–14].

Family-based and twin studies have proven a strong heritable component for both depression and obesity, with heritability estimates of ~35% and ~40% for depression and body mass index (BMI) respectively [15–17]. Although in both cases rare genetic variants and other chromosomal aberrations represent the bulk of the genetic load, genome-wide association studies (GWAS) have also identified a great number of associated single nucleotide polymorphisms (SNPs). These SNPs only represent a small fraction of the genetic susceptibility to these diseases nonetheless. While GWAS studies on BMI already identified hundreds of associated SNPs more than a decade ago [18–21], GWAS performed on depression have had notable difficulties for identifying associated variants [22]. Indeed, it has not been until quite recently that two depression GWAS meta-analyses identified 44 [23] and 102 [24] independent and significant loci associated with the disorder. Besides each individual genetic characterization, a shared genetic susceptibility profile has also been revealed for both conditions, which could be another influencing factor for the bidirectional depression–obesity relationship. Particularly, it has been estimated that up to 12% of the genetic component for depression could be shared with obesity [3, 25]. This promising finding has led to innovative approaches aiming to unveil the molecular mechanisms underlying the depression–obesity relationship [26–29].

Although initial expectations for GWASs on depression were high, mentioned SNPs individually account for only small proportions of reported heritability. Consequently, the practice of utilizing individual SNPs to predict depression is now considered a limited approach and other innovative perspectives have emerged to take advantage of available GWAS insights [30]. On this matter, several genomic studies have proposed to study multiple common SNPs collectively to improve the estimation of disease predisposition [31]. Based on the construction of genetic-risk scores (GRSs), that include multiple genetic variants at the same time, these approaches have recently gathered considerable interest [32], and have proven utility identifying groups of individuals who could benefit from the knowledge of their probabilistic susceptibility to disease. In brief, a GRS is usually calculated as a sum of the number of risk alleles carried by an individual, where the risk alleles are defined by the SNPs and their measured effects as detected by GWAS in a particular trait [33]. Although some authors have previously evaluated the performance of GRSs to discriminate depression [23], no study to date has investigated the utility of GRSs for depression prediction in people with obesity accounting for BMI information. On this matter, and given the strong relationship between obesity and depression, it could be possible that the inclusion of BMI information into the model (along with the GRS) elicits an improvement in predictive ability. Previous results from our group have already proved the hypothesis but in the opposite direction; a GRS for obesity improved its performance when the model included information about the depression status of each patient [34].

Therefore, in the present study we aimed: (i) To investigate whether a GRS combining a number of well-defined SNPs associated with depression might have utility for depression prediction in individuals with obesity, (ii) To evaluate whether the predictive ability of the model improves when obesity information is considered as a covariate, and (iii) To obtain a general picture of the cellular and molecular pathways mapped by those SNPs included into the GRS.

METHODS

Study population

The sample consisted of 2123 community-based individuals (136 depression cases, 1987 controls) from the PredictD-CCRT study; a Cluster, Controlled, Randomized Trial (CCRT). The PredictD-CCRT study was a national, multicentre randomized controlled trial, which had two parallel groups: cluster assignment by primary care center, and a follow-up of 18 months. The aim of this study was to assess the performance of a preventive intervention on the depression incidence, taking into account the level and profile of risk of depression of each individual. The PredictD-CCRT was conducted in 70 primary care centers from 7 Spanish cities. The Spanish National Health Service covers over 95% of the population, providing free medical service, which ensured a representative sample from the south of Spain. Participants were assessed for clinical, psychological, sociodemographic, anthropometric, lifestyle, and other environmental variables. Individuals who also agreed to participate in genetic studies gave specific informed consent and provided a biological sample. This study was approved in each participating city by the corresponding ethics committee, and it was conducted in compliance with the Helsinki Declaration. The PredictD protocol, effectiveness, and cost-effectiveness analyses are fully described and available elsewhere [35–38]. Briefly, patients belonging to the recruiting centers were selected using a systematic random sampling, each 4–6 patients, from the family physician's appointment lists at random starting points for each day. Family physicians further checked whether the selected patients met any of the following exclusion criteria: age under 18 or over 75 years; inability to understand or speak Spanish; severe mental disorder (psychosis, bipolar, personality disorder…); cognitive impairment; terminal illness; the patient is scheduled to be out of the city more than four months during the 18 months of the follow-up; and persons who attend the primary care center on behalf of the person that initially has the appointment [35].

Characterization of depression

The psychiatric interview section was conducted by trained interviewers, independently from physicians. These research assistants completed a 20-h training by accredited instructors, in order to guarantee standardization. The section depression of the Composite International Diagnostic Interview (CIDI) was used for the assessment of depression. The CIDI [39, 40] which is a structured interview was used for the diagnosis of depression according to the DSM-IV criteria.

Characterization of BMI and obesity

Height and weight data from each individual were used to calculate body mass index (BMI) using the formula: weight in kilograms divided by height

in square meters (kg/m^2). International cut-off reference points were applied for obesity categorization (BMI < 25: normal weight, BMI ≥ 25: overweight, and BMI > 30: obesity). Underweight individuals (BMI ≤ 18.5) were excluded from analyses.

SNP selection

An extensive review of the literature was performed by the research team. Medline and Scopus databases were explored using relevant terms in the field of depression-associated genes (e.g., "depressive disorder", "major depressive disorder", "major depression", "MDD", "candidate gene", "SNP", "polymorphism", "loci"). SNPs were initially selected based on two criteria: (i) SNPs from candidate genes reported in case-control studies on depression and replicated in more than one independent study, or in loci having a significant potential role in depression (i.e., loci involved in well-established pathways associated with depression: the hypothalamic–pituitary–adrenal (HPA) axis [41] and the serotonergic system [42]) (n = 25); (ii) SNPs associated with depression from GWAS or meta-analyses establishing a P-value cut-off of $P \leq 7 \times 10^{-6}$ (n = 47). The information obtained from each approach was then combined and compared to the list of SNPs available from Illumina technology (San Diego, California), so that a definitive list of candidate variants was obtained: 6 and 10 SNPs initially selected from candidate gene studies and GWAS, respectively, were discarded in this step. Finally, 56 SNPs were selected for downstream analyses: (i) 19 SNPs from candidate gene association studies [43–55] and (ii) 37 SNPs associated with depression in GWAS or meta-analyses [56–68]. Detailed information for the final 56 candidate SNPs included in the analyses and their mapped loci is available in Supplementary Table 1a, b.

We further investigated functional and biological databases for additional information about identified top genetic regions on depression. Among the genes mapped by the selected SNPs, there are genes associated with depression-related pathways [25, 69–74]. A pathway enrichment analysis (PEA) of those genes was performed with the gprofiler2 R package [75]. Pathways that were significantly overrepresented in the gene list, as compared to all known genes in the genome, were obtained from Reactome and WikiPathways databases (Supplementary Fig. 1). The P-value of the enrichment of the different pathways was computed using a Fisher's exact test and multiple-test corrections were applied, considering false discovery rates (FDRs), that measures expected proportion of false significant matches within results. Details for the enrichment analysis can be found in Supplementary Table 2. Besides, following a recent detailed protocol [76], statistically significant pathways (FDR-adj.P < 0.05) were used to build an enrichment map in the 'EnrichmentMap' Cytoscape app [77] with the aim to collapse redundant pathways into a single biological theme rather than including general and specific pathways with many shared genes (Supplementary Fig. 2).

Genotyping

A saliva sample was obtained from each participant using the Oragene DNA saliva collection kit (OG-500; DNA Genotek Inc.). DNA extraction was performed using standard procedures. DNA concentration was measured by absorbance measure using the Infinite® M200 PRO multimode reader (Tecan, Research Triangle Park, NC).

Genotyping was performed using the TaqMan® OpenArray^{TM} Genotyping System (Applied Biosystems, Foster City, CA) following the manufacturer's instructions. Raw data were analyzed with the TaqManGenotyper v1.2 software (Thermo Fisher Scientific). SNPs showing a linkage disequilibrium (LD) value of $R^2 > 0.8$ in pairwise unphased correlations were removed from the selection. For all candidate markers, we further evaluated call rate, Hardy–Weinberg equilibrium (HWE), and minor allele frequency (MAF). MAFs of all SNPs were ≥0.05 and similar to those reported for Iberian populations in Spain in phase 3 of the 1000 Genomes Project. To account for the presence of genotyping errors, all SNPs with less than a 95% call rate were excluded from the analyses. In relation to HWE, the Wigginton's exact test was applied only in controls at an alpha level of 0.05. After all quality control checks, the selected 56 markers were available for downstream analyses. A complete workflow detailing the whole SNP selection procedure can be found in Fig. 1.
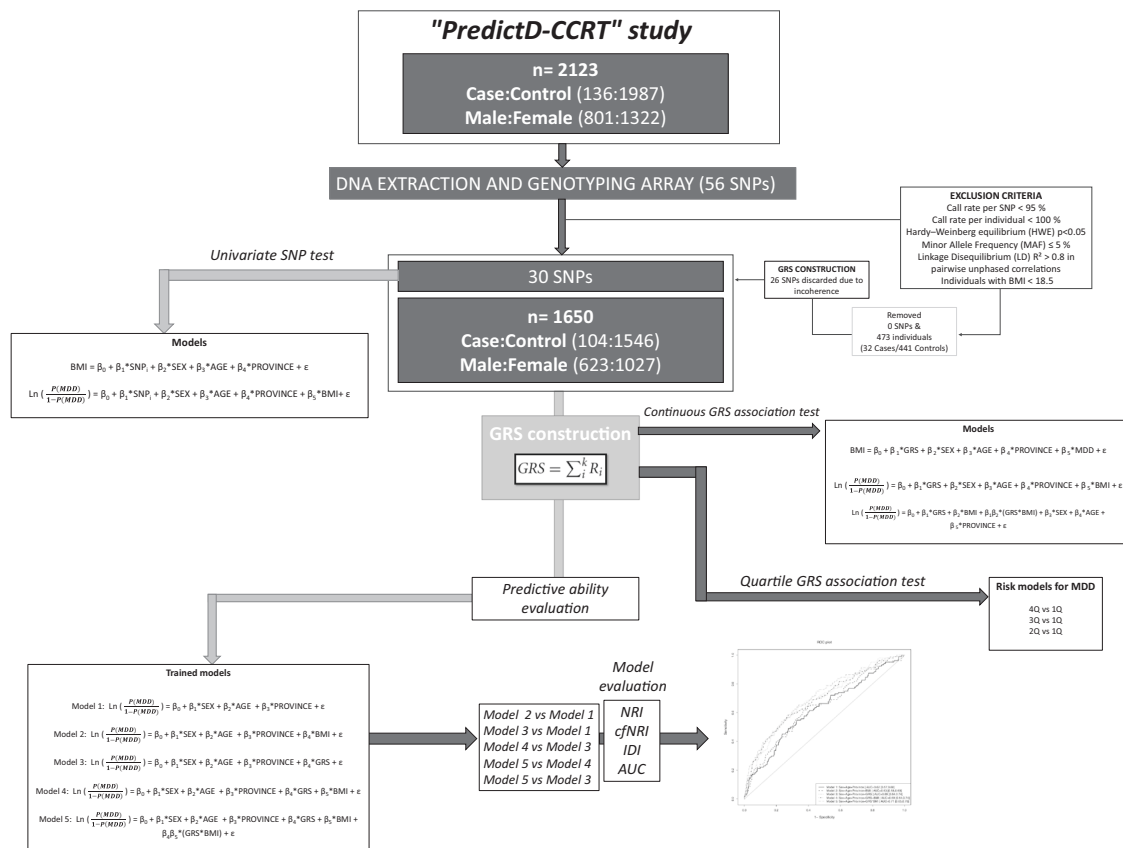
Fig. 1 Complete workflow detailing the study design and statistical analyses performed: quality control process, association analysis and construction and evaluation of predictive models. AUC area under the receiver operating characteristic curve, cfNRI the category-free net reclassification improvement, HWE Hardy–Weinberg equilibrium, IDI integrated discrimination improvement, LD linkage disequilibrium, MAF minor allele frequency, MDD major depressive disorder, NRI net reclassification improvement, SNP single nucleotide polymorphism.

Genetic-risk score construction (GRS)

To explore whether common variants with small risk effects on depression predict depression occurrence in our sample, a GRS was constructed following an unweighted approach, as implemented in the *PredictAbel* R package [78]. In this approach, a GRS is calculated for each individual based on the number of risk alleles for depression, without accounting for each SNP effect size. The motivation underlying this choice is the fact that most of the selected markers came from individual genetic studies, where no robust beta values were available. All initially selected SNPs from the literature (56 SNPs, as can be seen in Supplementary Table 1a), were assessed for association in our sample applying univariate regression analyses on depression. Although most of these SNPs were not significantly associated with depression in our sample, all effect sizes and direction of associations were compared to those reported in the literature for each selected SNP. With the aim of constructing a robust genetic-risk score, we only selected those SNPs showing concordant associations (in terms of effect size direction) between the association signal reported in the literature and the association signal reported in our sample. As a result, 30 SNPs from the 56 markers were selected for the construction of the GRS. The rationale behind this procedure is the fact that many of the selected SNPs might not be significant in our population while still being associated with depression according to other studies. For these SNPs that individually do not show statistical association, constructing a GRS is of special interest since we can combine their small effects so joined, they report a significant relationship with depression. SNPs in which the minor allele was reported as a protective marker (instead of a risk variant) were flipped in order to compute the risk score. Since the GRS cannot be estimated if any of the markers present a non-callable genotype for a certain individual, subjects showing a call rate <100% for any of the 30 candidate SNPs were removed from the analyses. As a result, 1650 individuals (104 depression cases and 1546 controls) were considered for the GRS construction. Details for the SNPs finally included in the GRS can be found in Supplementary Table 1a.

Statistical analysis

A complete workflow detailing the study design and statistical analyses performed can be found in Fig. 1. Differences between cases and controls for main clinical characteristics were analyzed using the Student's *t* test, the Welch's test, or the U Mann–Whitney test for quantitative variables. The Pearson's Chi-squared test was used for investigating group differences in categorical variables instead. Cohen's *d* and Cramér's *V* were calculated to assess effect sizes in quantitative and qualitative variables, respectively.

Binary logistic regression models were employed to test the effect of each individual SNP on depression under an additive genetic model of inheritance (thereafter named in the paper as univariate SNP analyses). We performed a post hoc power analysis based on a Z test for logistic regression using G-Power software. Power estimation was conceived for a logistic regression model including the GRS and the rest of adjusting covariates. Under the assumption of a normal distribution for the GRS in our sample, the power of our logistic regression in the $N = 1650$ sample, was estimated 99.99% (e.g., there is a 99.99% chance of correctly rejecting the null hypothesis that a particular value of the GRS is not associated with the value of the outcome variable (depression), with 1650 participants). Multiple linear regressions were used for the univariate SNP analyses on BMI. Regression models were evaluated by model control investigating linearity of effects on outcome(s), consistency with a normal distribution, and variance homogeneity. Continuous variables were tested for normality using the Shapiro–Wilk test and transformed when necessary by means of the natural log or the rank-based inverse normal transformation. All regression models employed are detailed in Fig. 1. Given the number of genetic markers analyzed, we considered false discovery rates (FDRs) calculated as in Benjamini and Hochberg to correct for multiple-hypothesis testing in univariate SNP analyses [79]. Regarding GRS analyses, logistic regression models were applied to test whether higher genetic-risk scores were observed for depression cases than controls. Logistic regression models were also applied for comparing participants presenting a high-risk genetic profile (Q2, Q3, or Q4) *vs*. those belonging to the reference quartile (Q1). Multiple linear regression was employed to investigate association between continuous GRS and BMI. Model deviance of logistic regressions (D²) was calculated to assess the amount of outcome variability explained by each group of variables. All tested models in our work were properly adjusted by confounders such as sex, age, province (geographical location), or BMI whenever necessary (Fig. 1).

To assess the predictive ability of the constructed GRS, five different predictive models were trained and evaluated in our sample (see trained models in Fig. 1). The area under the receiver operating characteristic (ROC) curve (AUC) was calculated for each model and all possible comparisons between constructed models were tested for statistical significance in terms of prediction improvement. Besides AUC, three recently proposed statistical metrics were also adopted to quantify the added predictive value of each model with respect to its immediate prior. These statistical metrics were the integrated discrimination improvement (IDI), the net reclassification improvement (NRI), and the category-free NRI (cfNRI). All of them have been previously described [80]. Since no established risk categories exist in depression, the NRI was applied according to standard risk categories (low (<5%), medium (5% to <25%), or high (≥25%)) [80]. For this reason, the cfNRI and the IDI were preferred estimators than NRI in our study. Discrimination plots, predictiveness curves, prior posterior risk curves, and risk distribution plots were obtained for each trained model (data not shown but available upon request). All predictive assessments were conducted using the *PredictABEL* and the *pROC* R packages [78, 81]. All statistical analyses were performed in R environment, version 3.4.5 (R Project for Statistical Computing).

Besides constructing our models in the whole population (interpretable models), we also implemented a sampling validation procedure. Particularly, K-fold cross-validation is a great method in case the classes are not equally balanced in a dataset. The use of this sort of validation consists of dividing the dataset in k groups or folds of samples (of equal sizes, if possible). Thus, the learning process is done with k−1 folds (training), and the evaluation of model's performance is done with the fold left out (testing). This iterative method helps to create models using different folds, and evaluate the model's performance through different metrics (with the fold left out). In Particular, we opted for a 5-fold cross-validation. To create the folds we used the function createFolds() from the caret package. To calculate the rest of the metrics we used the reclassification() function from the predictABLE package.

RESULTS

Demographic characteristics

A complete workflow detailing the adopted study design and the statistical analysis conducted can be found in Fig. 1. General characteristics of the study population by experimental condition are summarized in Supplementary Table 3. After excluding subjects with any missing genotypes or BMI under 18.5, a total of 1650 participants were finally included in our analysis (104 depression cases and 1546 controls). A significantly higher BMI was found in depression cases than in control ($P = 0.02$; Cohen's $d = 0.24$). Furthermore, we found statistically significant differences in the case-control proportion between the different provinces of recruitment ($P = 0.008$; Cramér's $V = 0.103$), suggesting that the geographical location of participants could be an important confounding variable to adjust genetics models for. The mean age of the participants was slightly higher in the group with depression than in controls, although this difference was not statistically significant ($P = 0.19$; Cohen's $d = 0.13$). Regarding sex, 62.24% of the total sample were females with no sex differences observed between cases and controls ($P = 0.64$; Cramér's $V = 0.12$).

Genetic association analyses on depression

Five SNPs from the 56 candidate genetic variants showed a significant association with depression status in the univariate analyses (Supplementary Table 4). While the reference minor alleles of SNPs rs6537837 and rs242939 (mapping the *GNAI3* and *CRHR1* genes, respectively) were reported as protective markers for depression, the reference minor alleles for the rs349475, rs310501, and rs1800532 (mapping the genes *LINC02223;CDH18*, *VCAN;HAPLN1*, and *TPH1*) were identified as risk markers for depression. Univariate SNP analyses for depression were adjusted for all pertinent confounders as illustrated in Fig. 1. Although none of these results remained statistically significant after strict multiple-hypothesis correction by FDR (alpha = 0.05), all associations were nominally confirmed using Fisher's exact tests. The genetic location of each SNP (i.e., exonic, intronic, intergenic, upstream, or downstream) is shown in Supplementary Table 1a.
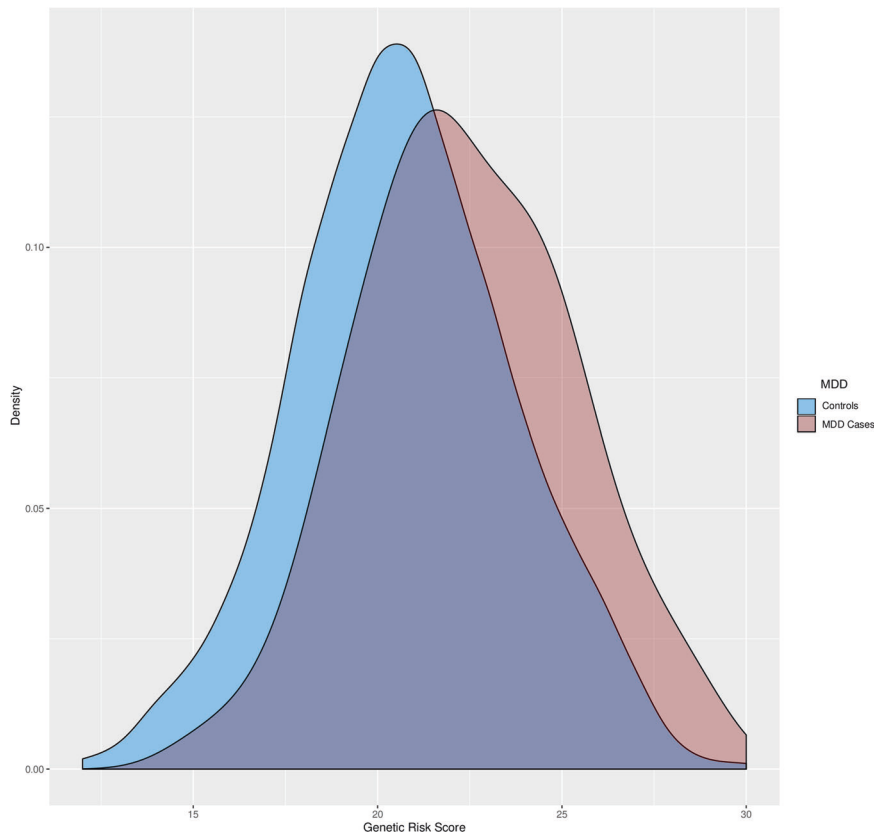
Fig. 2   Density distribution plot of the constructed GRS in our population. MDD major depressive disorder.

Surprisingly, all significant SNPs were identified as intronic or intergenic variants.

From these analyses, 30 SNPs were carefully selected (as they showed similar directional effect compared to the literature findings) and incorporated into an unweighted GRS. The GRS was tested for association with depression status as described in the "Methods" section and in Fig. 1. The density distribution plot of the constructed GRS in our population is presented in Fig. 2. The mean (and standard deviation) of the GRS in the whole sample was 20.82 (2.97), being 22.38 (2.91) in depression cases and 20.71 (2.94) in controls, being this difference statistically significant ($P = 1.02 \times 10^{-7}$; Cohen's $d = 0.57$) (Supplementary Table 3). Interestingly, a logistic regression model adjusted for sex, age, province, and BMI revealed a strong risk association between the GRS and the depression status, so that the odds of being depressed were estimated to increase by factor 1.35 for each additional risk allele in the GRS (OR = 1.35; CI 95% = [1.13, 1.27]; $P = 1.35 \times 10^{-7}$). The depression variability attributable to the genetic component in the model was estimated at 4.17%. When comparing individuals presenting the highest risk scores (Q4) to those belonging to the first quartile (Q1) a stronger association was evidenced (OR = 4.19; CI 95% = [2.3, 7.62]; $P = 2.76 \times 10^{-6}$). When comparing individuals in the third quartile (Q3) to those belonging to the first quartile (Q1), the association was quantified with an OR = 2.37 (CI 95% = [1.25, 4.46]; $P = 0.008$). The remaining comparison (Q2-*vs*-Q1) reported a non-significant result otherwise (OR = 1.73; CI 95% = [0.92, 3.26]; $P = 0.08$). When modeling BMI on depression in a model adjusted for sex, age and province, no significant association was reported (OR = 1.2; CI 95% = [0.98, 1.46]; $P = 0.09$). On the other hand, the most intriguing result was the interaction found between the GRS and BMI (further adjusted for age, sex, and province), with depression (OR = 1.14; CI 95% = [1.07–1.20]; $P = 1 \times 10^{-4}$). The direction and magnitude of this interaction can be observed in Fig. 3, and suggest the existence of

gene-environment interaction phenomena by which BMI increases the genetics-conferred risk of depression in high-susceptibility individuals.

Prediction of depression
To demonstrate the validity of the GRS for the prediction of depression, five logistic regression models were constructed, trained, and evaluated in our sample. These models comprised a model with classical demographic information only (age, sex, and province), named model 1, and other four additional models that further included (alone or combined) BMI or GRS information (see Fig. 1 for more details). The predictive ability of each model was evaluated using AUC. Statistical significance and magnitude of improvements in predictions between models were further assessed by means of the metrics NRI, cfNRI, and IDI, as described in the "Methods" section. Results for all models and performed comparisons are presented in Fig. 4 and Table 1. The lowest predictive ability corresponded to the model 1, incorporating classical demographic information only (AUC = 0.62, 95% CI = [0.57, 0.68]) (Fig. 4). The inclusion of the GRS into this model (model 3) reported an increase in the AUC up to 0.69 (95% CI = [0.64, 0.74]) and ($P = 1 \times 10^{-5}$ for cfNRI and IDI). Instead, the inclusion of BMI into the classical demographic model (model 2) did not provoke any improvement in the AUC of model 1 (Table 1). A model combining both the GRS and BMI as marginal terms alongside the classical demographic variables (model 4) barely improved the AUC reported for the model 3 (with the GRS as a marginal term). Surprisingly, a model incorporating an interaction term between the GRS and the BMI (model 5) achieved the higher predictive ability for depression (AUC = 0.71, 95% CI = [0.65, 0.76]). The significance of this improvement of model 5 with regard to both model 3 and 4 was estimated at $P = 0.009$ for IDI and NRI. Presented results correspond to the model constructed in the whole sample. In
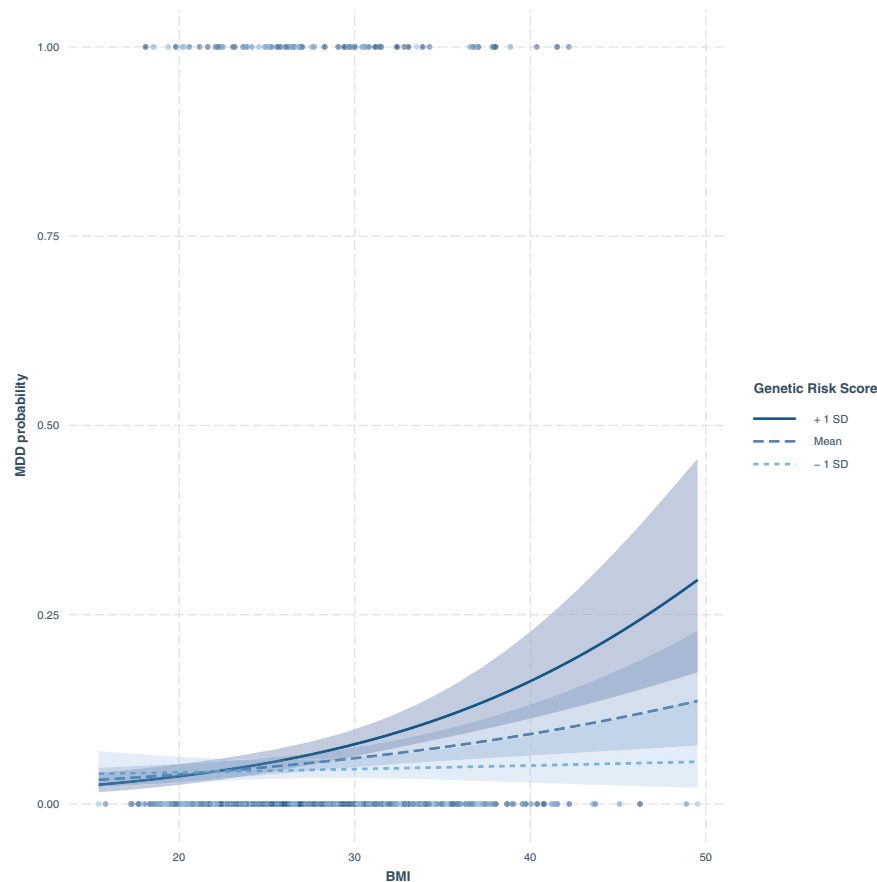
Fig. 3 Graphical representation of the direction and magnitude of the GRS*BMI interaction. BMI body mass index, MDD major depressive disorder, SD standard deviation.

addition, we cross-validated our findings employing a 5-fold CV procedure. Our main conclusions remained, although the AUC of all models slightly decreased (Supplementary Tables 5 and 6). The results shown in these tables correspond to the average metrics across folds (k = 5).

Genetic pleiotropy on BMI
Given previous evidence of a shared genetic-risk profile between depression and BMI, we investigated whether an association could exist between candidate SNPs and BMI in our sample. For that purpose, both univariate SNP and GRS-based analyses were performed with BMI as outcome variable (see Fig. 1 for more details regarding adjusting covariates). As a result, no significant association was reported between the GRS and BMI (OR = 1.2, 95% CI = [0.98−1.46], and $P = 0.09$) (Supplementary Fig. 3). In univariate SNP analyses, only the rs12457996 (mapping the *SYT4* gene) showed a significant association (Supplementary Table 7) being the C allele associated with a lower BMI in our sample ($P = 0.03$). The association did not remain statistically significant after multiple-testing correction.

DISCUSSION
In this study, we constructed an unweighted GRS, including 30 depression-associated genetic-risk variants from previous GWAS and candidate gene studies on depression (Fig. 1) [57–63]. As a first goal, we aimed to investigate whether the constructed GRS was associated with depression as well as if it was able to predict depression with enough precision and accuracy. Given the strong connection between depression and obesity, we also aimed to

elucidate whether the predictive ability of the GRS improved with the inclusion of BMI information for each individual. As a result, we found that the GRS is strongly associated with depression status and that it presents a not negligible depression predictive ability by itself. Remarkably, we showed how the interaction of BMI information with the GRS improves the predictive ability of the genetic component, deriving to a predictive ability of certain clinical relevance (AUC = 0.71). This result goes in line with recent approaches in which a close relationship between both conditions has been described [26–28] and complements our previous study in which we demonstrated the opposite relation [34].

We found that higher scores from the constructed GRS are strongly associated with a greater prevalence of depression in our sample ($P = 1.35 \times 10^{-7}$) (Fig. 2). In these analyses, the genetic component represented by the GRS accounted for 4.17% of the depression heritability. For the accomplishment of all these analyses, an unweighted GRS approach was employed, instead of a weighted GRS, due to the lack of GWAS or genetic meta-analyses providing robust estimates for the effects of the SNPs of interest in the literature [56–67]. Despite not using a weighted approach, our unweighted GRS demonstrated a strong association with depression, which is in line with previous reports on the clinical utility of GRS unweighted approaches [34, 82, 83].

The GRS showed a stronger association with depression than individual SNPs. Thus, it is quite probable that common genetic variants tested here represent only a small and cumulative contribution to the whole genetic susceptibility profile of depression [23, 24, 84], which is a commonly observed phenomenon in the genetic architecture of many complex diseases [85]. In our study, the finding of SNPs eliciting small
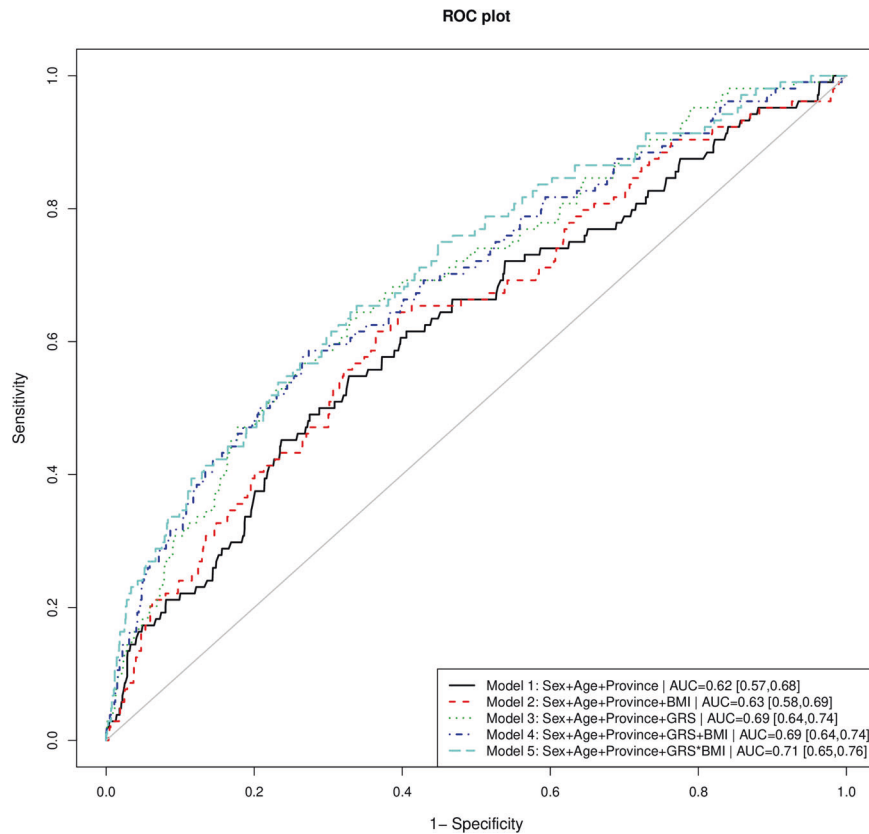
Fig. 4 Evaluation of the predictive ability of the constructed predictive model using AUC. AUC area under the receiver operating characteristic curve, GRS genetic-risk score.

Table 1. Statistics for model improvement with the addition of genetic and non-genetic-risk factors for MDD.

	Initial model: Model 1 Final model: Model 2	Initial model: Model 1 Final model: Model 3	Initial model: Model 3 Final model: Model 4	Initial model: Model 3 Final model: Model 5	Initial model: Model 4 Final model: Model 5
NRI	−0.03 (−0.08,0.009)	0.09 (−0.001,0.17)	−6e−04 (−0.04,0.04)	0.11 (0.02,0.19)	0.11 (0.02,0.19)
NRI *P*-value	0.12	0.05	0.98	**0.01**	**0.009**
cfNRI	0.19 (−0.006,0.39)	0.43 (0.24,0.63)	0.16 (−0.04,0.36)	0.24 (0.04,0.43)	0.17 (-0.03,0.37)
cfNRI *P*-value	0.06	**1e−05**	0.11	**0.01**	0.09
IDI	0.002 (−6e−04,0.005)	0.02 (0.01,0.03)	0.003 (5e−04,0.006)	0.02 (0.007,0.03)	0.01 (0.005,0.02)
IDI *P*-value	0.12	**1e−05**	**0.02**	**0.001**	**0.001**

Model 1 (Sex+Age+Province), Model 2 (Sex+Age+Province+BMI), Model 3 (Sex+Age+Province+GRS), Model 4 (Sex+Age+Province+GRS + BMI), and Model 5 (Sex+Age+Province+GRS*BMI). The 95% confidence intervals are shown in parentheses. Statistically significant results are highlighted in bold.
NRI net reclassification improvement, *cfNRI* category-free NRI, *IDI* integrated discrimination improvement, *AUC* area under the curve of the receiver operator characteristic curve.

and cumulative risk effects on depression was further supported by the fact that the greater and more significant differences were observed for the comparisons between individuals presenting a considerable number of risk variants (Q4) and individuals from the bottom reference quartile (Q1), which are individuals barely presenting risk alleles (OR = 4.19; CI 95% = [2.3, 7.62]; $P = 2.76 \times 10^{-6}$).

Both alone and in combination with classic demographic information, the GRS has demonstrated a good performance for the prediction of depression status in our sample. Contrary to the GRS, the addition of BMI information alone to the basic model did not show an improvement of its performance (although an increase in the AUC was reported, it did not reach statistical significance). The further inclusion of BMI information as an interaction term (along with the GRS) elicited a significant

improvement in the clinical prediction of depression (*P*-value IDI = 0.001 and *P*-value NRI = 0.009) (Table 1). Particularly, BMI was identified as a trigger-like risk factor for depression acting in a concerted way with the GRS component (Fig. 3). These findings, therefore, support the existence of a link between obesity and depression and reinforce the theory that the relationship is bidirectional.

In order to obtain a general picture of the cellular and molecular pathways mapped by analyzed SNPs, an enrichment functional analysis was conducted. Observed gene overlapping between some of the significant pathways suggests that those cellular mechanisms in which are involved could provide part of the disease map in depression. Associations between those obtained biological pathways and depression have been found in previous studies [25, 69, 70]. In this regard, for example, there is evidence

that confirms the antidepressant potential of ECM remodeling after chronic stress by mean of intra-cortical degradation of perineuronal nets (PNNs) [71]; or the role of the PI3K-Akt signaling pathway in proinflammatory regulatory mechanisms, given that neuron inflammation and inflammatory cytokine production contribute to the pathology of depression [72]; as well as, the alteration of circadian rhythms and disturbances of sleep [69, 73]. In this way, such evidence not only gives support to the use of certain polymorphisms as a predictive tool, but also helps us to contextualize which mechanisms are being altered and to have a more functional perspective of this analysis.

There are certainly some limitations that should be mentioned. Some of the main drawbacks from this study include a high unbalanced design between depression cases and control as well as the absence of analyzed SNPs from the recently published meta-analysis list [23, 24]. Therefore, generated hypotheses here would require more detailed characterization in bigger and independent cohorts.

In summary, we found that a GRS based on 30 depression-associated risk loci was significantly associated with depression. Although GRS on its own explained only a small amount of variance of depression, a significant novel feature of this study is that including non-genetic-risk factors such as BMI together with a GRS came close to the conventional threshold for clinical utility used in ROC analysis and improves the prediction of depression. This has potential clinical implications as well as implications for future research directions in exploring the links between depression and obesity-associated disorders. While it is likely that future genome-wide studies with very large samples will detect variants other than the common ones, it seems probable that a combination of non-genetic information will still be needed to optimize the prediction of obesity.

REFERENCES

1. Askari J, Hassanbeigi A, Khosravi HM, Malek M, Hassanbeigi D, Pourmovahed Z, et al. The relationship between obesity and depression. Procedia - Soc Behav Sci. 2013;84:796–800.
2. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 2006;3:e442.
3. Milaneschi Y, Simmons WK, van Rossum EFC, Penninx BW. Depression and obesity: evidence of shared biological mechanisms. Mol Psychiatry. 2019;24:18–33.
4. Farmer A, Korszun A, Owen MJ, Craddock N, Jones L, Jones I, et al. Medical disorders in people with recurrent depression. Br J Psychiatry. 2008;192:351–5.
5. Milaneschi Y, Lamers F, Berk M, Penninx BWJH. Depression heterogeneity and its biological underpinnings: toward immunometabolic depression. Biol Psychiatry. 2020;88:369–80.
6. Luppino FS, de Wit LM, Bouvy PF, Stijnen T, Cuijpers P, Penninx BWJH, et al. Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. Arch Gen Psychiatry. 2010;67:220–9.
7. Mannan M, Mamun A, Doi S, Clavarino A. Is there a bi-directional relationship between depression and obesity among adult men and women? Systematic review and bias-adjusted meta analysis. Asian J Psychiatr. 2016;21:51–66.
8. Mannan M, Mamun A, Doi S, Clavarino A. Prospective associations between depression and obesity for adolescent males and females—a systematic review and meta-analysis of longitudinal studies. PLoS ONE. 2016;11:e0157240.
9. Blaine B. Does depression cause obesity?: A meta-analysis of longitudinal studies of depression and weight control. J Health Psychol. 2008;13:1190–7.
10. Carpenter KM, Hasin DS, Allison DB, Faith MS. Relationships between obesity and DSM-IV major depressive disorder, suicide ideation, and suicide attempts: results from a general population study. Am J Public Health. 2000;90:251–7.
11. Stunkard AJ, Faith MS, Allison KC. Depression and obesity. Biol Psychiatry. 2003;54:330–7.
12. Penninx BWJH, Milaneschi Y, Lamers F, Vogelzangs N. Understanding the somatic consequences of depression: biological mechanisms and the role of depression symptom profile. BMC Med. 2013;11:129.
13. Schweinfurth N, Walter M, Borgwardt S, Lang UE. Depression and obesity. In: Ahmad SI, Imam SK, editors. Obesity. Cham: Springer International Publishing; 2016. p. 235–44.
14. Afari N, Noonan C, Goldberg J, Roy-Byrne P, Schur E, Golnari G, et al. Depression and obesity: do shared genes explain the relationship? Depress Anxiety 2010;27:799–806.
15. Ormel J, Hartman CA, Snieder H. The genetics of depression: successful genome-wide association studies introduce new challenges. Transl Psychiatry. 2019;9:114.
16. Robinson MR, English G, Moser G, Lloyd-Jones LR, Triplett MA, Zhu Z, et al. Genotype-covariate interaction effects and the heritability of adult body mass index. Nat Genet. 2017;49:1174–81.
17. Zaitlen N, Kraft P, Patterson N, Pasaniuc B, Bhatia G, Pollack S, et al. Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. PLoS Genet. 2013;9:e1003520.
18. Thorleifsson G, Walters GB, Gudbjartsson DF, Steinthorsdottir V, Sulem P, Helgadottir A, et al. Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. Nat Genet. 2009;41:18–24.
19. Willer CJ, Speliotes EK, Loos RJF, Li S, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet. 2009;41:25–34.
20. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet. 2010;42:937–48.
21. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015;518:197–206.
22. Levinson DF, Mostafavi S, Milaneschi Y, Rivera M, Ripke S, Wray NR, et al. Genetic studies of major depressive disorder: why are there no genome-wide association study findings and what can we do about it? Biol Psychiatry. 2014;76:510–2.
23. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat Genet. 2018;50:668–81.
24. Howard DM, Adams MJ, Clarke T-K, Hafferty JD, Gibson J, Shirali M, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nat Neurosci. 2019;22:343–52.
25. Amare AT, Schubert KO, Klingler-Hoffmann M, Cohen-Woods S, Baune BT. The genetic overlap between mood disorders and cardiometabolic diseases: a systematic review of genome wide and candidate gene studies. Transl Psychiatry. 2017;7:e1007.
26. Tyrrell J, Mulugeta A, Wood AR, Zhou A, Beaumont RN, Tuke MA, et al. Using genetics to understand the causal influence of higher BMI on depression. Int J Epidemiol. 2019;48:834–48.
27. Mulugeta A, Zhou A, Vimaleswaran KS, Dickson C, Hyppönen E. Depression increases the genetic susceptibility to high body mass index: evidence from UK Biobank. Depress Anxiety. 2019;36:1154–62.
28. Speed MS, Jefsen OH, Børglum AD, Speed D, Østergaard SD. Investigating the association between body fat and depression via Mendelian randomization. Transl Psychiatry. 2019;9:184.
29. Avinun R, Hariri AR. A polygenic score for body mass index is associated with depressive symptoms via early life stress: evidence for gene-environment correlation. J Psychiatr Res. 2019;118:9–13.
30. Schrodi SJ, Mukherjee S, Shan Y, Tromp G, Sninsky JJ, Callear AP, et al. Genetic-based prediction of disease traits: prediction is very difficult, especially about the future. Front Genet. 2014;5:162.
31. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet. 2018;50:1219–24.
32. Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. Nat Rev Genet. 2018;19:581–90.
33. Chatterjee N, Shi J, García-Closas M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. Nat Rev Genet. 2016;17:392–406.
34. Hung C-F, Breen G, Czamara D, Corre T, Wolf C, Kloiber S, et al. A genetic risk score combining 32 SNPs is associated with body mass index and improves obesity prediction in people with major depressive disorder. BMC Med. 2015;13:86.
35. Bellón JÁ, Conejo-Cerón S, Moreno-Peral P, King M, Nazareth I, Martín-Pérez C, et al. Preventing the onset of major depression based on the level and profile of risk of primary care attendees: protocol of a cluster randomised trial (the predictD-CCRT study). BMC Psychiatry. 2013;13:171.
36. Bellón JÁ, Conejo-Cerón S, Moreno-Peral P, King M, Nazareth I, Martín-Pérez C, et al. Intervention to prevent major depression in primary care: a cluster randomized trial. Ann Intern Med. 2016;164:656.
37. Fernández A, Mendive JM, Conejo-Cerón S, Moreno-Peral P, King M, Nazareth I, et al. A personalized intervention to prevent depression in primary care: cost-effectiveness study nested into a clustered randomized trial. BMC Med. 2018;16:28.
38. Moreno-Peral P, Conejo-Cerón S, de Dios Luna J, King M, Nazareth I, Martín-Pérez C, et al. Use of a personalised depression intervention in primary care to prevent

anxiety: a secondary study of a cluster randomised trial. Br J Gen Pr. 2021;71:
e95–104.

39. Rubio-Stipec M, Bravo M, Canino G. The Composite International Diagnostic Interview (CIDI): an epidemiologic instrument suitable for using in conjunction with different diagnostic systems in different cultures. Acta Psiquiatr Psicol Am Lat. 1991;37:191–204.

40. Robins LN, Wing J, Wittchen HU, Helzer JE, Babor TF, Burke J, et al. The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. Arch Gen Psychiatry. 1988;45:1069–77.

41. Pariante CM, Lightman SL. The HPA axis in major depression: classical theories and new developments. Trends Neurosci. 2008;31:464–8.

42. Köhler S, Cierpinsky K, Kronenberg G, Adli M. The serotonergic system in the neurobiology of depression: relevance for novel antidepressants. J Psychopharmacol. 2016;30:13–22.

43. Wu Y-L, Ding X-X, Sun Y-H, Yang H-Y, Chen J, Zhao X, et al. Association between MTHFR C677T polymorphism and depression: an updated meta-analysis of 26 studies. Prog Neuro-Psychopharmacol Biol Psychiatry. 2013;46:78–85.

44. Arias B, Fabbri C, Gressier F, Serretti A, Mitjans M, Gastó C, et al. TPH1, MAOA, serotonin receptor 2A and 2C genes in citalopram response: possible effect in melancholic and psychotic depression. Neuropsychobiology 2013;67:41–7.

45. Ching-López A, Cervilla J, Rivera M, Molina E, McKenney K, Ruiz-Perez I, et al. Epidemiological support for genetic variability at hypothalamic-pituitary-adrenal axis and serotonergic system as risk factors for major depression. Neuropsychiatr Dis Treat. 2015;11:2743–54.

46. Prestes AP, Marques FZC, Hutz MH, Bau CHD. The GNB3 C825T polymorphism and depression among subjects with alcohol dependence. J Neural Transm (Vienna). 2007;114:469–72.

47. Samaan Z, Anand SS, Anand S, Zhang X, Desai D, Rivera M, et al. The protective effect of the obesity-associated rs9939609 A variant in fat mass- and obesity-associated gene on depression. Mol Psychiatry. 2013;18:1281–6.

48. Jin C, Xu W, Yuan J, Wang G, Cheng Z. Meta-analysis of association between the -1438A/G (rs6311) polymorphism of the serotonin 2A receptor gene and major depressive disorder. Neurological Res. 2013;35:7–14.

49. Li X-B, Wang J, Xu A-D, Huang J-M, Meng L-Q, Huang R-Y, et al. Apolipoprotein E polymorphisms increase the risk of post-stroke depression. Neural Regen Res. 2016;11:1790–6.

50. López-León S, Janssens ACJW, González-Zuloeta Ladd AM, Del-Favero J, Claes SJ, Oostra BA, et al. Meta-analyses of genetic studies on major depressive disorder. Mol Psychiatry. 2008;13:772–85.

51. Kishi T, Yoshimura R, Fukuo Y, Okochi T, Matsunaga S, Umene-Nakano W, et al. The serotonin 1A receptor gene confer susceptibility to mood disorders: results from an extended meta-analysis of patients with major depression and bipolar disorder. Eur Arch Psychiatry Clin Neurosci. 2013;263:105–18.

52. Binder EB, Owens MJ, Liu W, Deveau TC, Rush AJ, Trivedi MH, et al. Association of polymorphisms in genes regulating the corticotropin-releasing factor system with antidepressant treatment response. Arch Gen Psychiatry. 2010;67:369–79.

53. Xu Z, Zhang Z, Shi Y, Pu M, Yuan Y, Zhang X, et al. Influence and interaction of genetic polymorphisms in the serotonin system and life stress on antidepressant drug response. J Psychopharmacol. 2012;26:349–59.

54. Papiol S, Arias B, Gastó C, Gutiérrez B, Catalán R, Fañanás L. Genetic variability at HPA axis in major depression and clinical response to antidepressant treatment. J Affect Disord. 2007;104:83–90.

55. Schumacher J, Jamra RA, Becker T, Ohlraun S, Klopp N, Binder EB, et al. Evidence for a relationship between genetic variants at the brain-derived neurotrophic factor (BDNF) locus and major depression. Biol Psychiatry. 2005;58:307–14.

56. Huang J, Perlis RH, Lee PH, Rush AJ, Fava M, Sachs GS, et al. Cross-disorder genomewide analysis of schizophrenia, bipolar disorder, and depression. AJP 2010;167:1254–63.

57. Sullivan PF, de Geus EJC, Willemsen G, James MR, Smit JH, Zandbelt T, et al. Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. Mol Psychiatry. 2009;14:359–75.

58. Lewis CM, Ng MY, Butler AW, Cohen-Woods S, Uher R, Pirlo K, et al. Genome-wide association study of major recurrent depression in the U.K. population. AJP. 2010;167:949–57.

59. Muglia P, Tozzi F, Galwey NW, Francks C, Upmanyu R, Kong XQ, et al. Genome-wide association study of recurrent major depressive disorder in two European case-control cohorts. Mol Psychiatry. 2010;15:589–601.

60. Rietschel M, Mattheisen M, Frank J, Treutlein J, Degenhardt F, Breuer R, et al. Genome-wide association-, replication-, and neuroimaging study implicates HOMER1 in the etiology of major depression. Biol Psychiatry. 2010;68:578–85.

61. Kohli MA, Lucae S, Saemann PG, Schmidt MV, Demirkan A, Hek K, et al. The neuronal transporter gene SLC6A15 confers risk to major depression. Neuron 2011;70:252–65.

62. Shi J, Potash JB, Knowles JA, Weissman MM, Coryell W, Scheftner WA, et al. Genome-wide association study of recurrent early-onset major depressive disorder. Mol Psychiatry. 2011;16:193–201.

63. Shyn SI, Shi J, Kraft JB, Potash JB, Knowles JA, Weissman MM, et al. 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. 2011;16:202–15.

64. Wray NR, Pergadia ML, Blackwood DHR, Penninx BWJH, Gordon SD, Nyholt DR, et al. Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. Mol Psychiatry. 2012;17:36–48.

65. Terracciano A, Tanaka T, Sutin AR, Sanna S, Deiana B, Lai S, et al. Genome-wide association scan of trait depression. Biol Psychiatry. 2010;68:811–7.

66. Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, et al. A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry. 2013;18:497–511.

67. CONVERGE consortium. Sparse whole-genome sequencing identifies two loci for major depressive disorder. Nature. 2015;523:588–91.

68. Aragam N, Wang K-S, Pan Y. Genome-wide association analysis of gender differences in major depressive disorder in the Netherlands NESDA and NTR population-based samples. J Affect Disord. 2011;133:516–21.

69. Walker WH, Walton JC, DeVries AC, Nelson RJ. Circadian rhythm disruption and mental health. Transl Psychiatry. 2020;10:28.

70. Chan KL, Cathomas F, Russo SJ. Central and peripheral inflammation link metabolic syndrome and major depressive disorder. Physiology. 2019;34:123–33.

71. Spijker S, Koskinen M-K, Riga D. Incubation of depression: ECM assembly and parvalbumin interneurons after stress. Neurosci Biobehav Rev. 2020;118:65–79.

72. Kitagishi Y, Kobayashi M, Kikuta K, Matsuda S. Roles of PI3K/AKT/GSK3/mTOR pathway in cell signaling of mental illnesses. Depression Res Treat. 2012;2012:1–8.

73. Li QS, Tian C, Seabrook GR, Drevets WC, Narayan VA. Analysis of 23andMe antidepressant efficacy survey data: implication of circadian rhythm and neuroplasticity in bupropion response. Transl Psychiatry. 2016;6:e889.

74. Fabbri C, Crisafulli C, Gurwitz D, Stingl J, Calati R, Albani D, et al. Neuronal cell adhesion genes and antidepressant response in three independent samples. Pharmacogenomics J. 2015;15:538–48.

75. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. 2019;47:W191–8.

76. Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, et al. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. Nat Protoc. 2019;14:482–517.

77. Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. PLoS ONE. 2010;5:e13984.

78. Kundu S, Aulchenko YS, van Duijn CM, Janssens ACJW. PredictABEL: an R package for the assessment of risk prediction models. Eur J Epidemiol. 2011;26:261–4.

79. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc: Ser B (Methodol). 1995;57:289–300.

80. Pickering JW, Endre ZH. New metrics for assessing diagnostic potential of candidate biomarkers. Clin J Am Soc Nephrol. 2012;7:1355–64.

81. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinforma. 2011;12:77.

82. Labos C, Thanassoulis G. Genetic risk prediction for primary and secondary prevention of atherosclerotic cardiovascular disease: an update. Curr Cardiol Rep. 2018;20:36.

83. Thanassoulis G, Peloso GM, Pencina MJ, Hoffmann U, Fox CS, Cupples LA, et al. A genetic risk score is associated with incident cardiovascular disease and coronary artery calcium: the Framingham Heart Study. Circ Cardiovasc Genet. 2012;5:113–21.

84. Sullivan PF, Agrawal A, Bulik CM, Andreassen OA, Børglum AD, Breen G, et al. Psychiatric genomics: an update and an agenda. Am J Psychiatry. 2018;175:15–27.

85. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 Years of GWAS discovery: biology, function, and translation. Am J Hum Genet. 2017;101:5–22.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

MR, AAR, EM, BG, and JC participated in the design of the study. MR led the study and supervised the analyses. BG, MR, JC, and EM were involved in the SNP selection process. BG and EM contributed to the coordination and genotyping of the samples. JC helped with the phenotypic and clinical data. AAR performed statistical analyses and wrote the first draft of the manuscript. AAR, JAZR, and ATM performed the construction of the genetic-risk score and subsequent statistical analyses. AMPG and ELI helped with the statistical analyses and performed the pathway enrichment analysis. JAB designed and coordinated the field work of the predictD-CCRT randomized controlled trial at the national level. PMP, SCC, JMA, MIBR, ARB, AF, CFA, CMP, and CMF co-designed the predictD-CCRT randomized controlled trial and contributed to the coordination of field work in their respective cities (Málaga, Bilbao, Jaén, Barcelona, Valladolid, Granada, and Zaragoza). PMP and SCC also contributed to the coordination at the National level. All authors discussed the results, provided critical feedback, and approved the submission.

All authors were involved in drafting the manuscript or revising it critically for important intellectual content, and approved the final version of the manuscript to be published.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41398-022-01783-7.

Correspondence and requests for materials should be addressed to Juan Antonio Zarza-Rebollo.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.