Supplementary material

Contents

1958 British Birth Cohort

The 1958 British Birth cohort (1958BC) includes all births during one week in March in England, Scotland and Wales [1]. Participants in the study have been followed regularly, with information collected on a wide-range of factors related to health, lifestyle, growth and development. At age 45 years, 11,971 participants currently living in Britain were invited to take part in a biomedical survey, of whom 9,377 (78%) filled in a questionnaire and 8,302 (89%) also provided a blood sample. The 1958BC is largely a white European population (98%) and despite some data attrition, it has been evaluated to be broadly representative of the surviving cohort [2]. Written consent was obtained from participants for the use of information in medical studies. The 45-year biomedical survey and genetic studies were approved by the South-East Multi- Centre Research Ethics Committee (ref: 01/1/44) and the joint UCL/UCLH Committees on the Ethics of Human Research (Ref: 08/H0714/40).

Genotyping

Genetic information was obtained from blood samples collected at 45y, through two sub-samples from case-control studies that had used the 1958BC as a source for population controls: 3000 samples were randomly selected as part of the Wellcome Trust Case Control Consortium [WTCCC2, [3]] and 2592 distinct samples were selected as part of the Type 1 Diabetes Genetics Consortium [T1DGC [4]]. The SNP rs16969968 was imputed both in T1DGC and in WTCCC2, with average posterior call rate >0.99 in both studies.

Depression and anxiety

At age 45 years, the Clinical Interview Schedule-Revised (CIS-R) was used to assess midlife common mental disorders [5]. This was administered by trained survey nurses during visits to cohort members' homes. A shortened form of the CIS-R was used for the 1958BC, in which sections regarding worry, obsessions, somatic symptoms, compulsions and physical health worries were not included, therefore focussing on depressive and anxiety disorders. Definitions of depression, anxiety and psychological distress were based upon 90th percentile cut offs of the summary scores. For depression, the summary score was the sum of 5 symptom scores (fatigue, forgetfulness, sleep problems, depression and depressive ideas) which ranges from 0-21. The anxiety summary score, is based on only one symptom score (anxiety) which ranges from 0-4. The continuous measures of depression, anxiety and psychological distress were based on the depression summary score, anxiety summary score and total CIS-R score (ranging from 0-37) respectively.

Smoking status

Cigarette smoking was recorded at age 42 by Computer Aided Personal Interviewing, and was classified as never, ex or current smoker. Reports of never smoking were verified using data from surveys at ages 23 and 33. Number of cigarettes smoked per day at 42 years was also reported for current smokers. Pipe and cigar smokers were excluded from analyses.

In the current analyses, we included 5,033 participants with rs1051730 genotype, depression and smoking data available.

ALSPAC

The Avon Longitudinal Survey of Parents and Children is a prospective cohort study which recruited pregnant women residing in Avon, United Kingdom, with expected dates of delivery between 1 April 1991 and 31 December 1992. Full details of the study recruitment and methodology have been published previously [6]. A total of 14,541 pregnancies were included in the initial sample, resulting in 14,062 live births and 13,988 children who were alive at one year of age. Detailed information on mothers and their partners (during and after pregnancy) and the children (since birth) has been collected from self-report questionnaires and attendance at clinics. Ethics approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committee.

Please note that the study website contains details of all the data that is available through a fully searchable data dictionary (http://www.bristol.ac.uk/alspac/researchers/data-access/data-dictionary/).

ALSPAC mothers

Genotyping

Genotyping of rs1051730 in the ALSPAC mothers was carried out by KBiosciences (Hoddesdon, UK; www.kbioscience.co.uk) using fluorescence-based competitive allele-specific PCR (KASPar). Full details of the genotyping methods have been published previously [7]. There was little evidence for deviation from Hardy Weinberg Equilibrium for this SNP in the ALSPAC mothers [7].

Depression and anxiety

Women completed the Edinburgh Postnatal Depression Scale (EPDS) in a postal questionnaire administered when their offspring were 3 years old. This questionnaire has 10 questions, each with a maximum score of 3 points, with a higher score indicating a greater level of depressed mood. For case definition 1, a cut-off point of greater than or equal to 13 on the EPDS was used to classify women as depressed. This is a well-established cut-off point for determining low mood and has been used previously in ALSPAC [8 9].

In the same questionnaire, women answered the Crown Crisp Experiential Index (CCEI) [10]. Scores on the CCEI anxiety subscale can range from 0 to 16. Women who scored 9 or higher, a cut point which has been used previously in ALSPAC [11], were classified as having anxiety according to case definition 1.

Smoking status

At the same time as the anxiety and depression measures, women were asked about current cigarette smoking status and daily number of cigarettes smoked. From this information, women were classified as current smokers if they smoked more than one cigarette per day. Women were classified as former and never smokers according to information provided in a questionnaire administered at 18 weeks gestation, in which they were asked about pre-pregnancy smoking habits. Women who only reported smoking cigars and pipes were excluded from analyses.

In the current analyses, we included 5,181 participants with rs1051730 genotype, depression and smoking data available.

ALSPAC children

Sample

A single individual from each family was used in the analysis. For twins, only the first born twin contributed to the analysis sample. Where there were multiple births by the same mother within the study, a single child was selected at random from each family.

Genotyping

DNA was extracted as described previously [12]. Genotyping was undertaken by KBioscience Ltd. (www.kbioscience.co.uk), using a proprietary competitive allele specific PCR system (KASPar) for SNP analysis.

Depression and anxiety

Depression and anxiety were measured via a computerised questionnaire, the Clinical Interview Schedule- Revised (CIS-R) [13], administered at a clinic which the ALSPAC children attended at age 17 years. The CIS-R that establishes the severity of 14 symptoms which constitute anxiety and depression disorders using algorithms based on ICD-10 criteria and has been validated against clinical diagnoses by trained psychiatrists [13]. Depression was defined as meeting ICD-10 criteria for a mild, moderate or severe episode of unipolar depression (case definition 1) or being above the 90th percentile of the depression symptom scale (the sum of the depression, fatigue, concentration, sleep and depressive thoughts symptom scores) (for case definition 2). Anxiety was defined as a primary or secondary diagnosis of anxiety (for case definition 1) or being above the 90th percentile of the sum of the anxiety (anxiety and worry) symptoms scores (for case definition 2). The depression symptom scale, anxiety symptoms score and total CIS-R score were used for continuous measures of depression, anxiety and psychological distress respectively.

Smoking status

At the age 17 clinic, participants also completed questions on a computer about their lifetime smoking behaviour. From these, two categories of smoking status were created: never smokers and current daily smokers. Never smokers reported never having tried a cigarette in their lifetime and current daily smokers smoked at least one cigarette per day. Individuals reporting less frequent smoking were excluded from analyses.

In the current analyses, we included 1,643 participants with rs1051730 genotype, depression and smoking data available.

BRHS

The British Regional Heart Study (BRHS) recruited 7,735 men aged 40–59 years in 1978–80; full details are reported elsewhere [14]. Men were recruited from 24 medium-sized British towns; at that time, very few eligible subjects were of non-European ancestry. Twenty years later, when aged 60–79 years, *n* = 4,252 participants were re-measured. The BRHS has local (from each of the districts in which the study was based) and multi-centre ethical committee approvals. All men provided informed written consent to the investigation, which was performed in accordance with the Declaration of Helsinki.

Genotyping

DNA was extracted from a whole-blood sample taken at the twenty year follow up. Genotyping of rs1051730 was performed by KBioscience (http://www.kbioscience.co.uk), using KASPar chemistry, which is a competitive allele-specific PCR SNP genotyping system.

Depression

Depression data was collected in a self-completed postal questionnaire in November 1996, 16-18 years after entry to the study (2-4 years before the twenty year follow up). Participants were asked "Have you ever been told by a doctor that you have or have had any of the following conditions?", with "depression" being included in a list of 11 medical conditions..

Smoking status

In the 1996 questionnaire, at the same time data on depression were collected, men were asked "do you currently smoke cigarettes?" and "if no, have you ever smoked cigarettes?", thus defining whether participants could be classed as never, ex- or current smokers. Men were also asked "how many cigarettes per day do you smoke?"

In the current analyses, we included 2,943 participants with rs1051730 genotype, depression and smoking data available.

BWHHS

The British Women's Heart and Health Study (BWHHS) is a prospective cohort study of women aged between 60 and 79 years who were randomly selected from general practice lists in 23 towns in England, Scotland, and Wales [15]. Between 1999 and 2001 baseline data were collected from 4,286 women with interviews by a research nurse, physical examination, review of primary care medical records and self-completion questionnaire. Follow-up for disease outcomes is by biennial medical record review (with validation checks) and cancer registrations and death certificates obtained from the Health and Social Care Information Centre. In 2003, 2007 and 2010 follow-up data were collected using self-reported questionnaires. Ethical approval was granted for the BWHHS from the London Multi-Centre Research Ethics Committee and 23 Local Research Ethics Committees.

Genotyping

DNA was extracted from blood samples collected in 1999-2001. Genotyping of rs1051730 was performed by KBioscience (http://www.kbioscience.co.uk), using KASPar chemistry, which is a competitive allele-specific PCR SNP genotyping system.

Depression

At baseline data collection, women self-reported lifetime diagnosis of depression by answering the question "Have you ever been told by a doctor that you have or have had any of the following conditions?", with "depression" being included in a list of 11 medical conditions.

Smoking status

In the baseline questionnaire, at the same time data on depression were collected, women were asked "have you ever smoked cigarettes regularly (at least 1/day)?" and "if yes, do you smoke cigarettes at present?", thus defining whether participants could be classed as never, ex- or current smokers. Women were also asked "how many cigarettes per day do you smoke?". Smoking heaviness was recorded as a continuous variable. Exclusive pipe and cigar smokers were excluded from the analyses.

In the current analyses, we included 3,150 participants with rs1051730 genotype, psychological distress, anxiety and smoking data available.

CaPS

The Caerphilly Prospective Study (CaPS) comprises 2,512 men—89% of all men aged between 45 and 59 years at baseline data collection (1979–1983) living in Caerphilly and adjoining villages [16]. The men have been followed an additional 4 times. During phase 2, an additional 447 men of the same age group living in the same area entered the study for the first time.

Ethics approval was obtained from the MRC Epidemiology Unit (Cardiff), the South Glamorgan Area Health Authority, the Gwent REC, and the South Wales Research Ethics Committee D, and each subject signed their agreement to be involved.

Genotyping

Genotyping of rs16969968 was undertaken by KBioscience Ltd. (www.kbioscience.co.uk), using a proprietary competitive allele specific PCR system (KASPar) for SNP analysis. The call rate was 98%, with 100% duplicate concordance.

Psychological distress and anxiety

At phase 2, men self completed the 30 item General Health Questionnaire (GHQ-30) which has been validated as a screening tool for common mental health disorders [17]. Respondents say how often they have experienced a symptom or behaviour in the past few weeks on a four point scoring system compared to usual (more than usual, same as usual, less than usual, much less than usual)[18].

The GHQ was used as a measure of psychological distress within this study. For binary outcomes, the GHQ was scored using the conventional scoring system which allocates 0,0,1,1 to the four response categories. The maximum score is 30 and high scores represent greater levels of mood disorder. We used a cut off of 5 or above to indicate psychological distress (for case definition 1); this cut point is commonly used in the literature [18 19]. When assessing the GHQ-30 as a continuous measure, we used the Likert scoring method to allow greater variation across the sample. The Likert method allocates scores of 0,1,2,3 to the responses, with a maximum score of 90 [17].

Anxiety at phase 2 was assessed using the 20 item State Trait Anxiety Inventory (STAI). Responses are on a four point scale (between 1 and 4), giving a maximum score (representing greater anxiety) of 80 [20]. No clear cut off points exist for this scale in the literature so we only applied a 90th percentile cut off (case definition 2). The scale score was used as a continuous measure of anxiety in the linear regression analysis.

Smoking status

Information on smoking (never, ex, and current smoker of 1–14, 15–24, or ≥25 cigarettes per day was collected by questionnaire at phase 2. Men who reported smoking cigars or pipes were excluded from all analyses.

In the current analyses, we included 1,113 participants with rs1051730 genotype, psychological distress, anxiety and smoking data available.

CHDS

The Christchurch Health and Development Study (CHDS) is a longitudinal study of a birth cohort of 1,265 children born in the Christchurch (New Zealand) urban region during a 4-month period in mid-1977 [21 22]. The cohort has now been studied on a total of 22 occasions from birth to age 30 years, using a combination of methods including: interviews with parents and cohort members; psychometric testing; teacher reports; biological sampling; and official record data. Rates of sample retention have remained high throughout the study and at age 30 a total of 987 participants (80% of the surviving cohort) were assessed. All phases of the study have received ethical approval from the regional Health and Disability Ethics Committee and all forms of data collection have been subject to the signed consent of research participants.

Genotyping

DNA samples were collected from participants between ages 28-30 years, via direct venous sampling (91.4% of samples), or Oragene™ saliva samples (8.6%). Genomic DNA was prepared from peripheral blood samples using a guanidium chloride extraction method followed by isopropanol precipitation [23]. Oragene™ samples were prepared by the method recommended by the supplier. DNA was resuspended in 10 mM Tris pH 8.0:1 mM EDTA (0.5 ml). Variation at rs16969968 was determined using a functionally tested 5' -nuclease TaqMan (Applied Biosystems) assay, on a 7500 Fast Real-Time PCR System (Applied Biosystems). For each reaction (5ul volume) approximately 50ng of genomic DNA was used, in ABsolute^{TM} Fast QPCR Master Mix, Low ROX (ABGene). Thermal cycling conditions were 15 minutes at 95°C, followed by 40 cycles of 95°C for 3 seconds and 60°C for 30 seconds. For quality assurance of the genotyping, 136 samples were repeated blind and no mismatches were identified. The resulting genotypes were tested and found to be in Hardy-Weinberg equilibrium.

Depression and anxiety

At age 30 participants were interviewed about aspects of their mental health and adjustment since the previous assessment. As part of this interview information relevant components of the Composite International Diagnostic Interview (CIDI) [24] were used to assess DSM-IV symptom criteria for a range of mental disorders. For the purposes of this analysis depression was defined as having had an episode of major depression in the previous 12 months. Anxiety was defined as having had an anxiety disorder (generalised anxiety disorder, panic disorder, agoraphobia, social phobia or specific phobia) in the previous 12 months.

Smoking status

Smoking status was classified on the basis of repeated assessments of cigarette smoking behaviour obtained at ages 15, 16, 18, 21, 25 and 30 years. At each assessment participants were questioned about their current smoking behaviour and the amount smoked. No questioning was conducted about other forms of tobacco use (pipes, cigars). For the purposes of this analysis current smokers were defined as those participants who reported smoking 1 or more cigarettes per day at age 30. Former smokers were defined as those who reported having smoked at least one cigarette per day at some time prior to age 30 but who were not current smokers. Never smokers were defined as

those who had never reported smoking at any age. Participants who reported some tobacco use but who had never become regular smokers (defined as at least one cigarette per day) were excluded from the analysis, as per study protocol. For current smokers smoking heaviness was defined on a 3-level categorical variable (1-9 cigs/day; 10-19 cigs/day; 20+ cigs/day).

In the current analyses, we included 697 participants with rs16969968 genotype, depression and anxiety and smoking data available.

CoLaus/PsyCoLaus

This is a population-based study of 6,188 individuals (CoLaus), aged 35–75 years, randomly selected from the list of residents in Lausanne, Switzerland, between 2003 and 2006 [25]. Risk factors for cardiovascular diseases were assessed and DNA and plasma samples were collected for the study of genetic variants and biomarkers. Between 2004 and 2008, all 35- to 66-year-old individuals of the CoLaus sample were invited to participate in an extensive psychiatric evaluation (PsyCoLaus) [26]. Detailed descriptions of recruitment procedures and assessments have been provided previously [25 26]. Colaus and PsyCoLaus were approved by the Institutional Ethic's Committee of the University of Lausanne.

Genotyping

Genotyping was performed on whole blood samples. The rs1051730 variant was genotyped as part of genome-wide SNP genotyping using the Affymetrix 500 K SNP chip. Full details of this have been described previously [7]. The SNP rs1051730 was imputed in the study sample with a posterior call rate of 0.97.

Depression and anxiety

Psychiatric assessment was made during the psychiatric evaluation (PsyCoLaus: 2004-08). Information on current diagnoses of depression and anxiety were collected using a French version of the semi-structured Diagnostic Interview for Genetic Studies (DIGS)[27 28]. Diagnoses of depression (major depressive disorder) and anxiety (generalised anxiety disorder, panic disorder, agoraphobia, social phobia or specific phobia) were assigned according to DSM-IV criteria. All interviews were conducted by master-level psychologists.

Smoking status

Smoking information was collected by self-rating questionnaires, approximately 1 year before the psychiatric evaluation.Ever smokers were those who had smoked regularly at some point in their life. Of these, individuals who smoked at least one cigarette daily at the time of interview were classified as current smokers, whereas the rest were classified as former smokers. Pipe and cigar smokers (who do not smoke cigarettes) were excluded. Smoking heaviness was collected as a continuous variable.

In the current analyses, we included 2,085 participants with rs1051730 genotype, depression, anxiety and smoking data available.

ELSA

The English Longitudinal Study of Ageing (ELSA) is a population based study of adults aged ≥50 years living in the UK. The original sample was recruited in 2002 and consisted of 11,391 individuals [29]. The sample has been followed up every 2 years and data have been collected via computer-assisted personal interviews and self-completion questionnaires. ELSA has been approved by the National Research Ethics Service and all participants have given informed consent.

Genotyping

In Wave 2 (2004/5) of the study, 5,633 participants provided blood samples for DNA extraction. Genotyping was performed by KBioscience using in-house KaSPAR technology.

Depression and anxiety

Current depression was assessed at Wave 1 (2002/3) using a shortened version of the Centre for Epidemiologic Studies Depression Scale (CES-D) [30]. The scale was formed of 8 questions each with a yes/no response and responses summed to give a scale ranging from 0-8. Depression was defined as a score of ≥4 (for case definition 1), a cut point which has been used and validated previously [31]. At the same wave, individuals were asked if they had ever been diagnosed by a doctor with anxiety.

Smoking status

Smoking status was defined according to information collected at Wave 1. Individuals were classified as current smokers if they reported smoking at least one cigarette per day or at least one gram of tobacco per day on weekdays. Former smokers were individuals who reported ever having smoked cigarettes but who were not current smokers. Never smokers were individuals who had never reported smoking cigarettes. Current or former pipe and cigar smokers who did not also smoke cigarettes were excluded from all analyses.

In the current analyses, we included 5,176 participants with rs1051730 genotype, depression, anxiety and smoking data available.

FINRISK

The National FINRISK Study is a large population survey on risk factors of chronic non-communicable diseases in Finland [32]. Every five years since 1972, area, sex and age stratified random samples of population have been drawn from the Population Information System. In these analyses data from FINRISK 2002 and 2007 surveys were used. In the 2002 and 2007 surveys, the age range of participants was from 25 to 74 years. Surveys have included a self-administered questionnaire, physical examination and blood draw for laboratory analyses and extraction of DNA. The 2002 and 2007 FINRISK surveys have been approved by the Coordinating Ethics Committee of the Helsinki University Hospital District. Each participant has given a written informed consent.

Genotyping

DNA was derived from whole blood samples, which were frozen immediately at the clinical study sites. The samples were transferred to the National Institute of Health and Welfare, where the DNA was extracted. Genotyping of rs16969968 (CHRNA5 D398N) was done under standard protocols of iPLEX Gold technology on the MassARRAY System (Sequenom, San Diego, CA, USA)[33]. The success rate was >0.99 and it was in HWE. Minor allele frequency was 0.32.

Depression

Depression was reported in a self-completed questionnaire. Cases were individuals who had had depression diagnosed or treated by a physician during the previous year.

Smoking status

In the same questionnaire, respondents were asked whether they had ever smoked. Those stating that they had never smoked were categorized as never smokers and skipped the other smoking-related questions. Ever smokers were defined as those who had smoked at least 100 cigarettes in their lifetime. Further questions were used to classify ever smokers as current and former smokers. Former smokers reported having been either regular or occasional smokers but were not smoking currently. For the present analysis, only those who had quit over 6 months ago were included in the former smoker category. Current smokers reported regular or daily smoking having smoked on the day of the assessment or the previous day. Exclusive pipe or cigar smokers were excluded from all analyses.

In order to create a variable for smoking quantity the participants were asked to indicate the average number of both manufactured and self-rolled cigarettes they smoked or had smoked per day before quitting. Manufactured and self-rolled cigarettes were totalled for the analysis.

In the current analyses, we included 10,566 participants with rs16969968 genotype, depression, and smoking data available.

Generation Scotland

The Generation Scotland: Scottish Family Health Study (GS:SFHS) recruited 24,084 participants aged 18–100 years between 2006–11; full details are reported elsewhere [34]. Participants came from across Scotland, with some family members from further afield. The sample was 59% female, with a wide range of ages and socio-demographic characteristics. Most (87%) participants were born in Scotland and 96% in the UK or Ireland. Mean family size (excluding 1,400 singletons without any relations in the study) was 4.07 members; median was 3 (IQR 2–5). The largest family had 36 participating members, and participants were grouped in 5,621 families. Although GS:SFHS is a family-based study, the current analysis only included subjects who were unrelated. Relatedness was defined using the pedigree data. An important feature of GS:SFHS is the breadth and depth of phenotype information, including clinical and physical measures and detailed data on cognitive function, personality traits and mental health. Ethical approval for the study was given by the NHS Tayside committee on research ethics (reference 05/s1401/89).

Genotyping

Genome-wide genotype data for nearly one million genetic variants have been measured on 10,000 selected participants. Blood samples (or saliva from postal and a few clinical participants) from GS:SFHS participants were collected, processed and stored using standard operating procedures and managed through a laboratory information management system at the Wellcome Trust Clinical Research Facility Genetics Core, Edinburgh [35]. The yield of DNA was measured using picogreen and normalised to 50ng/µl before genotyping. Genotyping was performed using the Illumina HumanOmniExpressExome-8 v1.0 DNA Analysis BeadChip and Infinium chemistry [36]. In summary, this consists of three steps: (i) whole genome amplification, (ii) fragmentation followed by hybridisation, and (iii) single-base extension and staining. For each of the samples, 4µl of DNA normalised to 50ng/µl was used. The Arrays were imaged on an Illumina HiScan platform and genotypes were called automatically using GenomeStudio Analysis software v2011.1.

Depression

Depression and smoking status were measured at the same time-point. Lifetime diagnosis of depression was assessed via the Structured Clinical Interview for DSM-IV Axis disorders non patient edition (SCIDI-NP)[37]. The interview was administered by a trained nurse.

The general health questionnaire 28 (GHQ-28) was administered at the same time point. For binary psychological distress, the GHQ-28 was scored using the conventional scoring system which allocates 0,0,1,1 to the four response categories. The maximum score is 28 and high scores represent greater levels of mood disorder. We used a cut off of 5 or above to indicate psychological distress (for case definition 1); this cut point has previously demonstrated good sensitivity and specificity for psychiatric disorders [38]. For the continuous measure of psychological distress, the GHQ-28 was scored according to the Likert method (a four point scale (0-3) according to symptom frequency) to allow greater variation across the sample with a range of 0-84 [17].

Smoking status

Smoking status was assessed via a paper-based, self-report questionnaire. Participants were asked a series of questions on smoking history, from which current smokers, former smokers and never smokers were defined. Pipe or cigar smokers who smoked no cigarettes were excluded from the analyses. Heaviness of smoking was assessed using the average number of cigarettes smoked per day.

In the current analyses, we included 7,485 participants with rs1051730 genotype, depression, and smoking data available.

GOYA females

The GOYA females were derived from the Danish genome-wide association study GOYA (Genomics of extremely Overweight Young Adults), nested within the Danish National Birth Cohort (DNBC) [39 40]. The Danish National Birth Cohort (DNBC) is a collection of data on 92,274 pregnant women recruited between 1996 and 2002, from their first antenatal visit to their general practitioner. Women participated in four telephone interviews (16 and 30 weeks gestation and 6 and 18 months after birth) and in a questionnaire-based follow-up 7 years after birth. They also provided two blood samples during pregnancy. The GOYA females were drawn as a random cohort sample from the 67,863 women within the DNBC who provided information about prepregnancy BMI, gave birth to a live born infant and provided a blood sample during pregnancy . This comprised 2,542 women. The study was approved by the regional scientific ethics committee and by the Danish Data Protection Board.

Genotyping

Genotyping of rs1051730 was carried out on the Illumina Human610-Quad v1.0 BeadChip (545,350 SNPs) at the Centre National de Génotypage (CNG), Evry, France.

Depression and anxiety

Information on treatment seeking for depression and anxiety was collected in the 7-year follow-up which had a participation rate of 55 %. Cases of depression and anxiety were defined as women who self-reported having seen a doctor for these conditions since birth using the following questions: 'Have you had a psychiatric illness or bad nerves since birth?', 'Have you been in contact with a doctor or psychologist because of this?' and 'Which type of illness were you suffering from?' with a positive answer for 'Anxiety' or ' Depression'.

Smoking status

Smoking status was based on information from the 7-year follow-up where women were asked whether they had smoked since birth, about their present smoking status (yes/no) and number of cigarettes smoked per day in three categories (1-9, 10-19, 20+). Past smokers were women who had quit smoking since birth and women who were non –smokers at the 7 year follow-up but were defined as smokers at any point in time during their pregnancy.

In the current analyses, we included 1,114 participants with rs1051730 genotype, depression, anxiety and smoking data available.

HBCS

The Helsinki Birth Cohort Study (HBCS) is composed of 8,760 individuals born between the years 1934-44 in one of the two main maternity hospitals in Helsinki, Finland. Between 2001 and 2003, a randomly selected sample of 928 males and 1,075 females participated in a clinical follow-up study with a focus on cardiovascular, metabolic and reproductive health, cognitive function and depressive symptoms. There were 1,709 women and men (43% men) with valid genotype and phenotype data. The mean age of the participants in the follow-up was 61.5 years (SD=2.9). In 2004, a psychological questionnaire was mailed to the participants, and in 2010-2011 a third follow-up was conducted. Detailed information on the selection of the HBCS participants and on the study design can be found elsewhere [41-43]. The research plan of the HBCS was approved by the Institutional Review Board of the National Public Health Institute and all participants have signed an informed consent.

Genotype

DNA was extracted from the blood samples and genotyping was performed with the modified Illumina 610k chip by the Wellcome Trust Sanger Institute, Cambridge, UK according to the standard protocols. rs1051730 was directly genotyped and did not deviate from the HWE ($p = 0.52$). Frequency of the A-allele was 34.0%.

Depression and anxiety

Information on depression and anxiety were collected by postal questionnaire in 2004 when individuals were 63 years old. Depression was assessed by the 20 item Centre for Epidemiologic Studies Depression Scale (CES-D) [30]. Each item on the scale is scored on a four point scale from 0-3 according to frequency of symptoms and summed to give a total score, with the maximum possible score (indicating the most severe depression) of 60. Individuals scoring 16 or above were classified as depressed, a cut point which has been applied previously in the literature [44 45].

Anxiety was assessed at the same time using the 20 item State Trait Anxiety Inventory (STAI) [20]. Responses are on a four point scale (between 1 and 4), giving a maximum score (representing greater anxiety) of 80 [20]. No clear cut off points exist for this scale in the literature so we only applied a 90^{th} percentile cut off (case definition 2).

Smoking status

As smoking information was not collected at the same time point as the depression and anxiety data (2004), smoking status from the initial follow up (2001 to 2004) was used. It was assumed that smoking status had not changed in the interim. Smoking status was self-reported on a following scale: never smoked, quitted smoking, smoke only on weekends, smoke 1-5 cigarettes/day, smoke 6-20 cigarettes /day, smoke >20 cigarettes /day. Current smokers were defined as individuals who reported smoking at least one cigarette per day. Data on pipe and cigar smoking were not available.

In the current analyses, we included 1,442 participants with rs1051730 genotype, depression, anxiety and smoking data available.

Health2006

The Health2006 study took place during 2006-2008 and consisted of a random sample of 7,931 Danish (Danish nationality and born in Denmark) men and women aged 18-69 years invited to participate in a health examination [46]. A total of 3,471 (43.8%) persons participated. Potential participants living in the Copenhagen area were identified in the central Danish Civil Registration System, and then recruited by invitation. The aim was to investigate the prevalence and risk factors of chronic diseases such as mental health, asthma, allergies, CVD, and diabetes. Informed written consent was obtained from all participants. The study was approved by the Ethical Committee of Copenhagen.

Genotyping

Blood samples were taken from all participants as part of their health examination. The buffy coat was frozen for DNA extraction, and later genomic DNA was extracted using a Qiagen AutoPure LS system. Genotyping was performed using KBiosciences allele-specific PCR (KASPar) (KBiosciences, Hoddesdon, UK). The call rate for this SNP (rs1051730) was > 99.2%. No errors were observed in 370 duplicate samples.

Depression and anxiety

Participants were asked to complete a self-administered shortened version of the Danish Symptom Checklist SCL-90-R [47] during the health examination. The shortened questionnaire included symptom subscales for depression (13 questions) and anxiety (10 questions). Each question is answered on a 5 point Likert scale, ranging from zero (not at all) to four (very much). The time frame is the past week. The scores are then summed for each subscale and divided by the number of items to give an average score for each scale of between 0 and 4. A higher score indicates more symptoms. As we were unable to identify any clear cut points for the scale in the literature, we used 90th percentile cut offs of the specific subscales to categorise individuals as depressed or anxious (case definition 2).

Smoking status

Information on smoking status was collected by a self-administered questionnaire to be filled in at home prior to the health examination. Smoking status was recorded as never, former, occasional (<1 cigarette, cheroot, cigar, or pipe per day) and daily smokers. Occasional smokers as well as daily smokers smoking exclusively cheroots, cigars, or pipe were excluded from all analyses. Smoking heaviness among daily smokers was recorded as number of cigarettes per day.

In the current analyses, we included 2,758 participants with rs1051730 genotype, depression, anxiety and smoking data available.

Health2008

The Health2008 Study took place from September 2008 to December 2009. Participants from the Copenhagen area were recruited through the Civil Registration office as a random sample of 2,218 men and women between 30 and 60 years of age [48 49]. Pregnant women, men and women with cardiovascular disease, type 2 diabetes, chronic obstructive lung disease or functional limitation were not eligible. The final study sample comprised 795 (35.8 %) eligible participants. The aim was to validate various methods applied in The Health2006 Study, e.g. a cardio respiratory fitness test and an IQ test. Informed written consent was obtained from all participants. The study was approved by the Ethical Committee of Copenhagen.

Genotyping

DNA was extracted from blood samples taken from all participants as part of their health examination. Genotyping was performed with the Illumina Human Exome BeadChip (version 1.2) using the Illumina HiScan (Illumina, San Diego, CA). Genotypes were called using the genotyping module (version 1.9.4) of GenomeStudio software (version 2011.1; Illumina). The call rate for this SNP (rs16969968) was > 99%.

Depression and anxiety

Participants were asked to complete a self-administered shortened version of the Danish Symptom Checklist SCL-90-R [47] during the health examination. The shortened questionnaire included symptom subscales for depression (13 questions) and anxiety (10 questions). Each question is answered on a 5 point Likert scale, ranging from zero (not at all) to four (very much). The time frame is the past week. The scores are then summed for each subscale and divided by the number of items to give an average score for each scale of between 0 and 4. A higher score indicates more symptoms. As we were unable to identify any clear cut points for the scale in the literature, we used 90th percentile cut offs of the specific subscales to categorise individuals as depressed or anxious (case definition 2).

Smoking status

Information on smoking status was collected by a self-administered questionnaire to be filled in at home prior to the health examination. Smoking status was recorded as never, former, occasional (<1 cigarette, cheroot, cigar, or pipe per day) and daily smokers. Occasional smokers as well as daily smokers smoking exclusively cheroots, cigars, or pipe were excluded from all analyses. Smoking heaviness among daily smokers was recorded as number of cigarettes per day.

In the current analyses, we included 585 participants with rs1051730 genotype, depression, anxiety and smoking data available.

HUNT

The second wave of the HUNT Study in Norway took place in 1995-97, where all adults aged 20 years and older in Nord Trøndelag County were invited to participate. A total of 65,438 (70%) accepted to the invitation and gave written informed consent to use the data for medical research. The data collection included questionnaires, clinical measurements and blood samples (http://www.ntnu.edu/hunt/data/que). Use of data in the present study was approved by the Regional Committee for Medical Research Ethics (Reference nr. 2013/1127/REK midt).

Genotyping

Altogether 56,664 participants were genotyped for the rs1051730 single nucleotide polymorphism variant. DNA was extracted from blood samples for all participants of the HUNT 2 study and stored at the HUNT biobank. The rs1051730 polymorphism was genotyped at the HUNT biobank using TaqMan genotyping assays (Applied Biosystems, Foster City, CA, USA) and performed on an Applied Biosystems 7900HT Fast real-Time PCR System using 10 ng of genomic DNA. The call rate cut-off was set to 90%, and the genotype frequencies were in agreement with HapMap data.

Depression and anxiety

Depression and anxiety were assessed using the self-completed Hospital Anxiety and Depression Scale (HADS) [50]. This scale is made up of two 14 item subscales, which assess depression and anxiety. Scores of 8 or above were used to classify individuals as having depression or anxiety; these cut points have been used previously within HUNT [51]. In addition, 90th percentile cut offs of the specific subscales were used to categorise individuals as depressed or anxious (case definition 2).

Information on use of depression and anxiety medication was collected among those responding >0 days on the questions: “If daily use of medication last 12 months; how many months have you used medicine for depression and/or sedatives.”

Smoking status

Smoking status was measured with self-completed questionnaire data with a categorical variable, and the participants were classified as never smokers, former smokers or current smokers. Current smokers were asked how many cigarettes they smoked per day. Data on smoking behaviour was reported in the same questionnaire as the anxiety and depression measures. Exclusive pipe and/or cigar smokers were excluded from the analyses.

In the current analyses, we included 53,289 participants with rs1051730 genotype, depression, anxiety and smoking data available.

Inter99

The Inter99 study is a randomised controlled trial (CT00289237, ClinicalTrials.gov) investigating the effects of lifestyle intervention on CVD (N=61,301)[52 53]. We used baseline data from a random subsample of 12,934 men and women aged approximately 30, 35, 40, 45, 50, 55, or 65 years invited to participate in a health examination during 1999-2001. Participants were living in the Copenhagen area and were identified in the central Danish Civil Registration System, and recruited by invitation. A total of 6,784 (52.5%) participated. A subsampe of 1,948 individuals examined in the period 15th February 2000 to 31st January 2001 were asked to complete a shortened version of the Danish SCL-90-R beforehand (1,837 accepted). Only participants with a Northern European origin (Denmark, Norway, Sweden, Iceland, and Faeroe Islands) were included (n=1,770). Informed written consent was obtained from all participants. The study was approved by the Ethical Committee of Copenhagen.

Genotyping

DNA was extracted from blood samples taken from all participants as part of their health examination. Genotyping was performed using KBiosciences allele-specific PCR (KASPar) (KBiosciences, Hoddesdon, UK). The call rate for this SNP (rs1051730) was > 98.8%. No errors were observed in 353 duplicate samples.

Depression and anxiety

One month prior to the baseline screening participants were asked to complete a self-administered shortened version of the Danish Symptom Checklist SCL-90-R [47]. The use of this scale within Inter99 has been described previously [54]. The shortened questionnaire included symptom subscales for depression (13 questions) and anxiety (10 questions). Each question is answered on a 5 point Likert scale, ranging from zero (not at all) to four (very much). The time frame is the past week. The scores are then summed for each subscale and divided by the number of items to give an average score for each scale of between 0 and 4. As we were unable to find any clear cut points for the scale in the literature, we used 90th percentile cut offs of the specific scales to categorise individuals as depressed or anxious (case definition 2).

Smoking status

Information on smoking status was collected by a self-administered questionnaire to be filled in at home prior to the health examination. Smoking status was recorded as never, former, occasional (<1 cigarette, cheroot, cigar, or pipe per day) and daily smokers. Occasional smokers as well as daily smokers smoking exclusively cheroots, cigars, or pipe were excluded from all analyses. Smoking heaviness among daily smokers was recorded as number of cigarettes per day.

In the current analyses, we included 1,493 participants with rs1051730 genotype, depression, anxiety and smoking data available.

NFBC1966

The Northern Finland Birth Cohort 1966 (NFBC1966) Study was initiated in 1965 by enrolling mothers living in the two Northernmost provinces of Finland, Oulu and Lapland, and with expected dates of delivery in 1966 [55]. Altogether 12,231 children were born into the cohort, 12,058 of them live-born. The original data have been supplemented by data collected by postal questionnaires at the ages of 1, 14, 31 and 46 years (on-going) and data from various hospital records and national register data. At the 31 year follow up, those living in the original target area (Northern Finland) or in the capital (Helsinki) area were invited to a clinical examination, in which 71% (N=6,033) participated. The University of Oulu Ethics Committee and the Ethical Committee of Northern Ostrobothnia Hospital District have approved the study. Participants provided written informed consent.

Genotyping

Blood samples were drawn at age 31 and DNA was extracted successfully for 5,753 subjects. Illumina's HumanCNV370-Duo DNA Analysis BeadChip was used to obtain genome-wide genetic data. It contains an informative set of tag SNPs derived from the HapMap European-derived (CEU) sample. Imputation was performed on 328,007 SNPs using IMPUTE software version 0.3.1 58, applying information threshold of >0.4 and MAF threshold of >1%. The rs1051730 SNP was extracted from GWAS data. The University of Oulu Ethics Committee and the Ethical Committee of Northern Ostrobothnia Hospital District have approved the study. Participants provided written informed consent.

Depression and anxiety

Depression and anxiety were assessed via a Finnish translation of the Symptom Checklist, SCL-25 which was administered in a postal questionnaire at age 31 years [56]. The SCL-25 questionnaire includes symptom subscales for depression (13 questions) and anxiety (10 questions). Each question is answered on a 4 point scale, ranging from one (not bothered) to four (extremely bothered). The scores are then summed for each subscale and divided by the number of items to give a total score for each scale of between 1 and 4. A cut off of 1.75 and above on each of the subscales was used to indicate depression or anxiety (for case definition 1), as previously described in the literature [57].

Smoking status

Smoking habits were enquired via a postal questionnaire as part of the 31-year follow-up study. The questionnaire was returned by 75% (N=8,767) of the individuals. Smoking heaviness was measured on a continuous scale. The exclusive pipe or cigar smokers excluded were excluded from all analyses.

In the current analyses, we included 3,919 participants with rs1051730 genotype, depression and smoking data available.

NFBC1986

The Northern Finland Birth Cohort 1986 (NFBC1986) includes all births in the two northernmost provinces of Finland, Oulu and Lapland, between July 1st 1985 and June 30th 1986 [58]. Altogether 9,479 children were born into the cohort, 9,432 of them live-born. The original data have been supplemented by data collected by postal questionnaires at the ages of 7, 8 and 16 years and data from various hospital records and national register data. A clinical examination was performed at age 16 years in which 74% (N=6,798) of the cohort members participated. The University of Oulu Ethics Committee and the Ethical Committee of Northern Ostrobothnia Hospital District have approved the study. Participants provided written informed consent.

Genotyping

Blood samples were collected as part of the 16-year follow-up and DNA was extracted for 6266 individuals. The DNA samples were processed at Imperial College London, UK and custom genotyping was performed at LGC Genomics Ltd, Hoddesdon, Herts, UK (formerly Kbioscience). The University of Oulu Ethics Committee and the Ethical Committee of Northern Ostrobothnia Hospital District have approved the study. Participants provided written informed consent.

Depression and anxiety

Depression and anxiety were assessed via the Youth Self Report (YSR), a psychiatric assessment tool for detecting emotional and behavioural problems in 11-18 year olds [59]. The YSR is self administered and adolescents respond on each statement on a scale of “not true”, “sometimes true” or “definitely true”. Scores on each question are added together to obtain a summary score for each scale. The YSR is made up of eight syndrome scales, two of which relate to anxiety and depression, the withdrawn/depressed (7 items) scale and the anxious/depressed (13 items) scale. Individuals scoring >84th percentile of the withdrawn/depressed subscale were classified as having depression and individuals scoring >84th percentile of the anxious/depressed scale were classified as having anxiety (for case definition 1). This is in line with previous cut points used to indicate borderline clinical range [59 60].

Smoking status

Smoking habits were enquired via a postal questionnaire as part of the 16-year follow-up study. The questionnaire was returned by 80% (N=7,344) of the adolescents. Two categories of smokers were created, never smokers who reported never having tried a cigarette and current smokers who reported smoking at least one cigarette daily. Individuals who did not fall into this category were excluded from the analyses.

In the current analyses, we included 1,143 participants with rs1051730 genotype, depression and smoking data available.

NHANES

The National Health and Nutrition Examination Survey (NHANES) (http://www.cdc.gov/nchs/nhanes.htm) is a program of health surveys run by the National Center for Health Statistics, part of the Centers for Disease Control and Prevention in the United States. The third National Health and Nutrition Examination Survey (NHANES III) was conducted from October 1988 through October 1994 in two phases, each of which comprised a national probability sample. The first phase was conducted from October 18, 1988, through October 24, 1991, at 44 locations. The second phase was conducted from September 20, 1991, through October 15, 1994, at 45 different locations. In NHANES III, 39,695 persons were selected over the six years; of those, 33,994 (86%) were interviewed in their homes. All interviewed persons were invited to the Medical examination centre (MEC) for a medical examination. Seventy-eight percent (30,818) of the selected persons were examined in the MEC, and an additional 493 persons were given a special, limited examination in their homes [61].

Data collection for NHANES was approved by the NCHS Research Ethics Review Board. Analysis of deidentified data from the survey is exempt from the federal regulations for the protection of human research participants. Analysis of restricted data through the NCHS Research Data Center is also approved by the NCHS ERB.

Genotype

The genetic data available through the NCHS is from 7,159 specimens collected during Phase-2 of the Third National Health and Nutrition Examination Survey. 2,630 specimens were collected from Non-Hispanic White individuals. Genotyping of rs1051730 was carried out using the iPLEX Gold (Sequenom, San Diego, CA) at Vanderbilt University 's DNA Resources Core. Average SNP call rate was 99.5%. This genotype is a restricted variable and therefore these data were accessed through the Research Data Center.

Depression

Depression was assessed at the MEC by the Diagnostic Interview Schedule (DIS) questionnaire which was administered by a trained interviewer. Individuals with lifetime diagnoses of a severe or non severe depressive episode (excluding those due to bereavement) according to DSM-III criteria were classified as depressed. If individuals were under the age of 16 years when they first experienced an episode of depression, they were excluded from the analyses.

Smoking status

Smoking status was collected in Household Adult Questionnaire administered during the household interview. Current smokers were defined as individuals who reported smoking at least one cigarette per day. Former smokers were defined as individuals who reported having smoked at least 100 cigarettes in their lifetime but did not currently smoke cigarettes. Never smokers were individuals who reported having smoked less than 100 cigarettes in their lifetime. Pipe and cigar smokers who did not also report cigarette smoking were excluded from analyses.

Sample analysis

Due to the survey design of the NHANES data, analyses were performed using survey commands available in Stata (version 11). Taylor series linearization was implemented to estimate variances [62]. Analyses were weighted using the genetic weights provided by NHANES, as described previously [63] . Analyses were restricted to individuals of non- Hispanic white ethnic origin.

In the current analyses, we included 694 participants with rs1051730 genotype, depression and smoking data available.

NSHD

The Medical Research Council National Survey of Health and Development (NSHD) is an on-going prospective birth cohort study consisting of all births in England, Scotland and Wales in one week in March 1946 [64]. The sample includes single, legitimate births whose fathers were in non-manual or agricultural occupations and a randomly selected one in four of all others, whose fathers were in manual labour. The original cohort, now 68 years of age, comprised 2,547 women and 2,815 men who have been followed-up over 20 times since their birth. The data collected to date include repeat cognitive function, physical, lifestyle and anthropomorphic measures, as well as blood analytes and other measures. The cohort recently completed a particularly intensive phase of clinical assessment and biological sampling with blood and urine sampling and analysis, and cardiac and vascular imaging [65]. Ethical approval was given by the Central Manchester Research Ethics Committee.

Genotyping

DNA was extracted from blood samples collected in 1999 [66]. Genotyping of rs16969968 was carried out by LGC Genomics (Hoddesdon, UK; www.lgcgenomics.com) using fluorescence-based competitive allele-specific PCR (KASPar). Call rate was >95%.

Depression and anxiety

The general health questionnaire 28 (GHQ-28) was used as a measure of psychological distress within this study and was administered to study members in their homes by trained interviewers at age 53 years [38 67]. This instrument has been validated for the identification of psychological distress [17]. Respondents say how often they have experienced a symptom or behaviour in the past few weeks on a four point scoring system compared to usual (more than usual, same as usual, less than usual, much less than usual)[18]. For binary outcomes, the GHQ was scored using the conventional scoring system which allocates 0,0,1,1 to the four response categories. The maximum score was 28 and high scores represent greater levels of mood disorder. We used a cut off of 5 or above to indicate psychological distress (for case definition 1); this cut point has previously demonstrated good sensitivity and specificity for psychiatric disorders [38]. For the continuous measure of psychological distress, the GHQ-28 was scored according to the Likert method (a four point scale (0-3)) to allow greater variation across the sample with a range of 0-84.

Smoking status

Smoking status was collected during a home interview at age 53 years by trained interviewers [68]. Current cigarette smoking status (“yes”, “no”) and the number of cigarettes smoked per day was obtained. Study members who provided an affirmative response to being current cigarette smokers, regardless of the quantity of cigarettes smoked per day, were classified as “smokers”, while those who provided a negative response were classified as “non-smoker”. Pipe and cigar smokers who did not also report cigarette smoking were excluded from analyses.

In the current analyses, we included 2,576 participants with rs16969968 genotype, depression and smoking data available.

NTR

Participants were registered with the Netherlands Twin Register (NTR). Longitudinal survey collection in twins and their parents started in 1991. Participants were invited about every 3 years to complete a survey containing questions about health, lifestyle, personality and psychopathology [69]. The data used in the present study derived from the 8th wave of survey collection, collected between 2009 and 2011. All NTR participants (twins and their family members) aged 18 years and older were invited to complete the survey. The NTR study was approved by the Central Ethics Committee on Research Involving Human Subjects of the VU University Medical Center, Amsterdam (IRB number IRB-2991 under Federalwide Assurance 3703; IRB/institute code 03-180), and all subjects provided written informed consent.

Genotyping

Genotyping at the rs16969968 locus was performed on different platforms from blood and/or buccal samples. Further details regarding DNA collection, genotyping, imputation and quality control checks in the NTR can be found elsewhere [70 71].

Whole blood and/or buccal DNA samples were collected as part of a large biobank project or other NTR projects [69 72]. DNA samples were genotyped in different projects and genotyping was performed on Affymetrix 6.0 (N=298), Affymetrix Perlegen 5.0 (N=3,697), Illumina 370 (N=290), Illumina 660 (N=1439), and Illumina Omni Express 1M (N=455) platforms. Calls were made with platform specific software (Genotyper, Beadstudio). The quality control thresholds for SNPs were MAF > 1%, HWE > 0.00001, call rate >95% and 0.30 <Heterozygosity <0.35. Samples were excluded from the data if their expected sex and IBD status did not match, or if the genotype missing rate was >10%. SNPs were aligned to the positive strand of the Hapmap-2-Build 36-release-24 CEU reference set. Alignment was checked using individuals and family members tested on multiple platforms. SNPs were excluded if allele frequencies differed more than 15% with the reference set and/or the other platforms. The data of the platforms were merged into a single dataset (N=5,856). This merged set was imputed against the reference set using IMPUTE v2. After imputation, genotype dosage was calculated if the highest genotype probability was above 90%. Badly imputed SNPs were removed based on HWE < 0.00001, proper info < 0.40, MAF < 1%, allele frequency difference >.15 against reference.

Depression and anxiety

The 8th NTR survey include the ASEBA Adult Self Report (ASR) questionnaire. Individuals with ASR T-scores ≥ 65 on the DSM-orientated depression scale were categorized as depressed, individuals with T-scores ≥ 65 on the the DSM-orientated anxiety scale were categorized as anxious (case definition 1)[73]. Furthermore, individuals were classified as having psychological distress if they met case definition 1 for either depression or anxiety. For case definition 2, individuals with scores on the ASR depression or anxiety scales above the 90th percentile were classified as depressed or anxious, and as psychologically distressed if they were above the 90th percentile on either the depression or anxiety scales.

Smoking status

Data on smoking behaviour were collected at the same time as the ASR, in survey 8 (2009) of the longitudinal survey study of the NTR. Subjects were asked whether they (ever) smoked and current smokers were asked how many cigarettes per day they smoked (continuous measure). Daily cigarette smokers were selected for all analyses, subjects who only smoked pipe or cigar smokers were excluded.

In the current analyses, we included 3,048 unrelated participants with genotypes at the rs16969968 locus for whom data on depression, anxiety and smoking were available.

Patch 2

The Patch 2 Study was the 8 year follow up to a pharmacogenetic study of smoking cessation with the nicotine transdermal patch. In the initial study, which took place between July 1999 and July 2000, 752 patients were recruited from general practice surgeries in Oxfordshire, UK [74]. In 2002, study participants were re-contacted and invited by letter to participate in a follow-up study to investigate relationships between genotype, personality and smoking behavior [75]. Ethical approval was obtained from the Anglia and Oxford Multicentre Research Ethics Committee and from the Local Research Ethics Committees covering the areas of residence of the patients.

Genotyping methods

Blood samples were collected at the 8 year follow up. DNA was extracted from plasma and buffy coat lymphocytes which were stored at –80°C until required for analysis. Genotyping of rs1051730 was performed has been described previously by primer extension and MALDI-TOF–based allele detection (Sequenom, San Diego, CA). Full genotyping methods have been described previously [76].

Depression

Depression was assessed using an adapted version of the depression section of the Structured Clinical Interview for DSM-III-R diagnosis (31), which was administered by self-report questionnaire. From this, a diagnosis of lifetime depression was obtained.

Smoking status

Smoking status was obtained by self-report at the time of depression assessment. As this population was recruited originally on the basis of being a current smoker, Patch 2 participants were either current or former smokers at the time of depression assessment.

In the current analyses, we included 232 participants with rs1051730 genotype, depression and smoking data available.

Rotterdam

The Rotterdam Study is an ongoing population-based cohort study of the elderly that started in 1990 and studies the incidence and determinants of chronic diseases in late life [77]. The Medical Ethics Committee of Erasmus Medical Center Rotterdam approved the study and a written informed consent was obtained from all participants .Every 3 to 4 years, all participants undergo an extensive home interview and a physical examination at the research center. In addition, all participants are continuously monitored for the occurrence of all major events during follow-up by linkage of the study database with medical files from general practitioners.

Genotyping methods

DNA was isolated from whole blood by manual salting-out protocol and was subsequently stored in Eppendorff tubes at -20 °C. DNA sample storage is in Matrix 2D-barcode tubes in 96 well format. Genotyping was based on 550 and 610 K Illumina arrays.

Depression and anxiety

Assessments for both depressive symptoms and anxiety disorders were performed in 2002-2004. Depressive symptoms were assessed using the validated Dutch version of the Center for Epidemiology Depression Scale (CES-D). A score of 16 or higher was considered suggestive of depressive symptoms [45]. An adapted version of the Munich version of the Composite International Diagnostic Interview (M-CIDI) [78]was administered to assess 1-year prevalence of the following anxiety disorders according to the DSM-IV-TR criteria [79]**:** generalized anxiety disorder, panic disorder, agoraphobia, social phobia, and specific phobia. For analysis, participants were categorized into two groups on basis of being free of any anxiety disorder or having at least one of the above mentioned anxiety disorders.

Smoking status

Information on smoking behaviour was obtained at the same time as depression and anxiety assessment using a computerized questionnaire during the home visit. Participants were classified as current smokers, former smokers, or never smokers. Current smokers were participants who answered yes to the question: "are you currently smoking?" Former smokers were participants who answered no to this question but who positively answered the question: "are you a former smoker?" Both current and former smokers answered additional questions on smoking habits (cigarettes, pipes, or cigars), amount of smoking, duration of smoking, and on age of starting and of stopping, if applicable. Exclusive pipe and cigar smokers were excluded from all analyses (very few in number).

In the current analyses, we included 6,520 participants with rs1051730 genotype, depression and smoking data available.

SYS-P

The Saguenay Youth Study (www.saguenay-youth-study.org) is a cross-sectional study investigating associations between prenatal exposure to maternal cigarette smoking and brain & behaviour and cardiovascular & metabolic health during adolescence [80]. Adolescents and their biological parents are recruited from a region with a known founder effect, namely the Saguenay Lac-Saint-Jean region of Quebec, Canada. Both the maternal and paternal grandparents of the adolescents are of French-Canadian ancestry born in the region; as such, all participants are of a single ethnicity.

We have acquired parental DNA and genotypes, and asked the parents – at the time of the adolescent recruitment - to fill out a number of questionnaires about themselves (e.g., mental health, substance use, employment, education). In the current meta-analysis, only data from the parents were included. The Research Ethics Committee of the Chicoutimi Hospital in Quebec, Canada approved the study protocol.

Genotyping

Genotyping of rs1051730 in the SYS-P parents was conducted at the Genome Analysis Centre of Helmholtz Zentrum Munchen (Munich, Germany) using the Illumina HumanOmniExpress BeadChip. Full details of the genotyping methods have been published previously [81].

Depression and anxiety

Participants completed a self-report questionnaire, developed for SYS by J. Séguin (Groupe de Recherche sur l'Inadaptation Psychosociale), consisted of the 12-item version of the Centre for Epidemiologic Studies Depression Scale (CES-D) [82]. Responses were on a 4-point scale (ranging from 1-4) with a total range of 12-48. For case definition 1, individuals scoring ≥24 were classified as depressed, which is equivalent to a previously used cut point on the 12 item version [82].

Within the same questionnaire, anxiety was assessed using a 10-item self-report question instrument, which covers each of the main DSM anxiety disorders. In a preliminary study, the inter-correlations between theoretically related pairs of DSM anxiety items were consistent (Séguin, Personal Communications). Participants were considered anxious according to case definition 1 if they scored either a 4 or a 5 (in a scale between 1-5) in at least two of the questions.

Smoking status

Women were asked about their current cigarette smoking status and daily number of cigarettes smoked in a telephone interview administered during recruitment of the families. The depression and anxiety measures were administered after the recruitment. Men were asked about their current cigarette smoking status and daily number of cigarettes smoked in the same self-report questionnaire developed for SYS-P by J. Séguin. This was questionnaire was administered at the same time as the depression and anxiety measures. Number of cigarettes per day was a continuous variable. Exclusive pipe or cigar smokers were excluded from all analyses.

Statistical analysis

As the SYS-P sample contains spouse pairs, possible clustering in the analysis was adjusted for by using robust standard errors. This was carried out using the “gee” package in R.

In the current analyses, we included 884 participants with rs1051730 genotype, depression, anxiety and smoking data available.

Whitehall II

Whitehall II recruitment of 10,308 participants (70% men) between 1985 and 1988 involved 20 London based Civil service departments [83]. Genetic samples were collected in 2004 from over 6,000 participants. The study is highly phenotyped for cardiovascular and other ageing related health outcomes, with 9 phases of follow up (5 with clinical assessment and biological sampling), over 20 years of follow up. A wide variety of health behaviour and environmental data are also collected and the participants are consented for linkage to recorded clinical data such as Hospital Episode Statistics (HES), the Office of National Statistics mortality data and the national registry of acute coronary syndromes in England and Wales (Myocardial Ischaemia National Audit Project).Ethical approval for the Whitehall II study was obtained from the University College London Medical School committee on the ethics of human research. Informed consent was gained from every participant.

Genotyping

Genotyping of rs16969968 was performed as part of genotyping using the Metabochip [84].

Depression and anxiety

At the first phase of the Whitehall II study, participants completed the 30 item General Health Questionnaire (GHQ-30) which has been validated as a screening tool for common mental health disorders [17]. Respondents say how often they have experienced a symptom or behaviour in the past few weeks on a four point scoring system compared to usual (more than usual, same as usual, less than usual, much less than usual)[18].

The GHQ was used as a measure of psychological distress within this study. For binary outcomes, the GHQ was scored using the conventional scoring system which allocates 0,0,1,1 to the four response categories. The maximum score is 30 and high scores represent greater levels of mood disorder. We used a cut off of 5 or above to indicate psychological distress (for case definition 1); this cut point is commonly used in the literature [18 19 85]. When assessing the GHQ-30 as a continuous measure, we used the Likert scoring method to allow greater variation across the sample. The Likert method allocates scores of 0,1,2,3 to the responses, with a maximum score of 90 [17].

Smoking status

Information on smoking status was collected by questionnaire during the first phase of data collection. Individuals who reported smoking cigars or pipes but not cigarettes were excluded from all analyses. Cigarettes per day was recorded as a continuous variable.

In the current analyses, we included 2,844 participants with rs16969968 genotype and psychological distress data available.

Supplementary Table S1 Case definitions for depression, anxiety and psychological distress in the CARTA studies

Cohort	Analysis	Scale	Diagnostic criteria	Scale Range	Case definition 1	Case definition 2
1958BC	Psychological distress	CIS-R	N/A		N/A	>90th percentile of depression OR anxiety
	Depression	CIS-R	N/A	0-21	N/A	>90th percentile of CIS-R depression symptom score
	Anxiety	CIS-R	N/A	0-4	N/A	>90th percentile of CIS-R anxiety symptom score
ALSPAC children	Psychological distress	CIS-R	ICD-10		Depression OR anxiety as characterised below	>90th percentile of depression OR anxiety
	Depression	CIS-R	ICD-10	0-21	Mild or moderate or severe depressive episode	>90th percentile of CIS-R depression symptom score
	Anxiety	CIS-R	ICD-10	0-8	Primary or secondary diagnosis of anxiety	>90th percentile of CIS-R anxiety symptom score
ALSPAC mothers	Psychological distress	EPDS/ CCEI	N/A	N/A	≥13 on EPDS OR ≥9 on CCEI	>90th percentile of EPDS or CCEI
	Depression	EPDS	N/A	0-30	≥13 on EPDS	>90th percentile of EPDS
	Anxiety	CCEI	N/A	0-16	≥9 on CCEI	>90th percentile of CCEI
BRHS	Psychological distress	N/A	N/A	N/A	N/A	N/A
	Depression	Doctor Diagnosis (lifetime)	N/A	N/A	Yes to doctor diagnosis	N/A
	Anxiety	N/A	N/A	N/A	N/A	N/A
BWHHS	Psychological	N/A	N/A	N/A	N/A	N/A

	distress					
	Depression	Doctor Diagnosis (lifetime)	N/A	N/A	Yes to doctor diagnosis	N/A
	Anxiety	N/A	N/A	N/A	N/A	N/A
CaPS	Psychological distress	GHQ-30	N/A	0-30	Overall GHQ-30 score ≥5 (conventional scoring)	>90th percentile of GHQ- (conventional scoring)
	Depression	N/A	N/A	N/A	N/A	N/A
	Anxiety	STAI (trait scale)	N/A	20-80	No established cut off	>90th percentile of STAI
CHDS	Psychological distress	CIDI	DSM-IV	N/A	Depression OR anxiety as characterised below	N/A
	Depression	CIDI	DSM-IV	N/A	Major depression in past 12 months	N/A
	Anxiety	CIDI	DSM-IV	N/A	Anxiety disorder (GAD, panic disorder, agoraphobia, social phobia, specific phobia) in past 12 months	N/A
CoLaus	Psychological distress	DIGS	DSM-IV	N/A	Depression OR anxiety as characterised below	N/A
	Depression	DIGS	DSM-IV	N/A	Major depression	N/A
	Anxiety	DIGS	DSM-IV	N/A	Anxiety disorder (GAD, panic disorder, agoraphobia, social phobia, specific phobia)	N/A
ELSA	Psychological distress	CES-D/Doctor diagnosis	N/A	N/A	Depression OR anxiety as characterised below	N/A
	Depression	CES-D (8 item)	N/A	0-8	≥4 on CES-D	>90th percentile of CES-D
	Anxiety	Doctor Diagnosis	N/A	N/A	Yes to doctor diagnosis	N/A

		(lifetime)				
FINRISK 2002/2007	Psychological distress		N/A	N/A	N/A	N/A
	Depression	Doctor diagnosis (previous 12 months)	N/A	N/A	Yes to doctor diagnosis	N/A
	Anxiety		N/A	N/A	N/A	N/A
Generation Scotland	Psychological distress	GHQ-28	N/A	0-28	≥5 using standard binary coding	>90th percentile on GHQ-28
	Depression	SCIDI-/NP	DSM-IV	N/A	Lifetime diagnosis of single major depressive disorder OR multiple major depressive disorder	N/A
	Anxiety	N/A	N/A	N/A	N/A	N/A
GOYA females	Psychological distress	Sought out doctor	N/A	N/A	Depression OR anxiety as characterised below	N/A
	Depression	Sought out doctor	N/A	N/A	Yes to depression	N/A
	Anxiety	Sought out doctor	N/A	N/A	Yes to anxiety	N/A
HBCS	Psychological distress	CES-D/STAI	N/A	N/A	≥16 on CES-D or >90th percentile of the STAI	>90th percentile of the CES-D or STAI
	Depression	CES-D	N/A	0-60	≥16 on CES-D	>90th percentile of CES-D
	Anxiety	STAI	N/A	20-80	No established cut off	>90th percentile of STAI
Health 2006	Psychological distress	SCL-90-R	N/A	N/A	N/A	>90th percentile on depression subscale OR anxiety subscale
	Depression	SCL-90-R	N/A	Average score from 13 items (0-4)	N/A	>90th percentile on depression subscale

	Anxiety	SCL-90-R	N/A	Average score from 10 items (0-4)	N/A	>90th percentile on anxiety subscale
Health 2008	Psychological distress	SCL-90-R	N/A	N/A	N/A	>90th percentile on depression subscale OR anxiety subscale
	Depression	SCL-90-R	N/A	Average score from 13 items (0-4)	N/A	>90th percentile on depression subscale
	Anxiety	SCL-90-R	N/A	Average score from 10 items (0-4)	N/A	>90th percentile on anxiety subscale
HUNT	Psychological distress	HADS	N/A	N/A	Depression OR anxiety as characterised below	>90th percentile on depression subscale OR anxiety subscale
	Depression	HADS	N/A	0-21 depression subscale	≥8 on depression subscale	>90th percentile on depression subscale
	Anxiety	HADS	N/A	0-21 anxiety subscale	≥8 on anxiety subscale	>90th percentile on anxiety subscale
Inter99	Psychological distress	SCL-90-R	N/A	N/A	N/A	>90th percentile on depression subscale OR anxiety subscale
	Depression	SCL-90-R	N/A	Average score from 13 items (0-4)	N/A	>90th percentile on depression subscale
	Anxiety	SCL-90-R	N/A	Average	N/A	>90th percentile on

				score from 10 items (0-4)		anxiety subscale
NFBC 1966	Psychological distress	SCL-25	N/A	N/A	Depression OR anxiety as characterised below	>90th percentile on depression subscale OR anxiety subscale
	Depression	SCL-25	N/A	Average score from 13 items (1-4)	≥1.75 depression subscale	>90th percentile on depression subscale
	Anxiety	SCL-25	N/A	Average score from 10 items (1-4)	≥ 1.75 anxiety subscale	>90th percentile on anxiety subscale
NFBC 1986	Psychological distress	YSR	N/A		Depression OR anxiety as characterised below	> 90th percentile of YSR depression or anxiety subscale
	Depression	YSR	N/A	(0-14) Withdrawn /depressed subscale (7 items)	≥84th percentile of depression subscale	> 90th percentile of YSR depression subscale
	Anxiety	YSR	N/A	(0-26) Anxious/ depressed subscale (13 items)	≥84th percentile of anxiety subscale	> 90th percentile of YSR anxiety subscale
NHANES	Psychological distress	N/A	N/A	N/A	N/A	N/A
	Depression	DIS	DSM III criteria	N/A	Lifetime single episode of major depression OR recurrent major	N/A

					depression	
	Anxiety	N/A	N/A	N/A	N/A	N/A
NSHD	Psychological distress	GHQ-28	N/A	0-28	≥5 using standard binary coding	>90th percentile on GHQ-28
	Depression	N/A	N/A	N/A	N/A	N/A
	Anxiety	N/A	N/A	N/A	N/A	N/A
NTR	Psychological distress	ASR	N/A			>90th percentile on depression subscale OR anxiety subscale
	Depression	ASR	N/A	0-28	T-scores ≥65 on depression subscale	>90th percentile on depression subscale
	Anxiety	ASR	N/A	0-14	T-scores ≥65 on anxiety subscale	>90th percentile on anxiety subscale
Patch 2	Psychological distress	N/A	N/A	N/A	N/A	N/A
	Depression	SCIDI	DSM-III-R criteria	N/A	Lifetime episode of major depression	N/A
	Anxiety	N/A	N/A	N/A	N/A	N/A
Rotterdam	Psychological distress	CES-D/diagnosis	N/A	N/A	Depression OR anxiety as characterised below	N/A
	Depression	CES-D (20 item)	N/A	0-60	≥16 on CES-D	N/A
	Anxiety	M-CIDI	DSM-IV	N/A	Anxiety disorder (GAD, panic disorder, agoraphobia, social phobia, specific phobia)	N/A
SYS-P	Psychological distress	CES-D/anxiety questionnaire	N/A	N/A	Depression OR anxiety as characterised below	>90th percentile on CES-D OR anxiety squestionnaire
	Depression	CES-D (12 item)	N/A	12-48	≥24 on CES-D	>90th percentile on CES-D
	Anxiety	10 item anxiety questionnaire	DSM	1-5	≥4 on at least 2 items of the anxiety questionnaire	>90th percentile on anxiety questionnaire

Whitehall II	Psychological distress	GHQ-30	N/A	0-30	≥5 using standard binary coding	>90th percentile on GHQ-30
	Depression	N/A	N/A	N/A	N/A	N/A
	Anxiety	N/A	N/A	N/A	N/A	N/A

CIS-R: Computerised interview schedule- revised, EPDS: Edinburgh Post-natal depression Scale, GHQ: General Health Questionnaire, STAI: State Trait Anxiety Inventory, CIDI: Composite International Diagnostic Interview, DIGS: Diagnostic Interview for Genetic Studies, CES-D: Centre for Epidemiologic Studies Depression, SCIDI/NP: Structured Clinical Interview for DSM-IV Axis disorders non patient edition, SCL: Symptoms checklist, HADS: Hospital Anxiety and depression scale, YSR: Youth Self Report, DIS Diagnostic Interview Schedule, ASR: Adult Self Report, SCID: Structured Clinical Interview for DSM-III-R diagnosis, M-CIDI: Munich version of Composite International Diagnostic Interview. All scales measure current depression and anxiety unless otherwise stated.Clinician diagnosis was assessed by self-reported recall in all studies.

Supplementary Table S2. Definitions of continuous measures of depression and anxiety in each study

Cohort	Analysis	Scale	Range	
1958BC	Psychological distress	CIS-R	0-37	Continuous scale of total CIS-R
	Depression	CIS-R	0-21	Continuous scale of depression symptoms
	Anxiety	CIS-R	0-4	Continuous scale of anxiety symptoms
ALSPAC children	Psychological distress	CIS-R	0-57	Continuous scale of total CIS-R
	Depression	CIS-R	0-21	Continuous scale of depression symptoms
	Anxiety	CIS-R	0-8	Continuous scale of anxiety symptoms
ALSPAC mothers	Psychological distress	N/A	N/A	N/A
	Depression	EPDS	0-30	EPDS continuous
	Anxiety	CCEI	0-16	CCEI continuous
CaPS	Psychological distress	GHQ-30	0-90	Continuous of GHQ-30 (Likert scoring)
	Depression	N/A	N/A	N/A
	Anxiety	STAI	20-80	Continuous STAI (trait)
ELSA	Psychological distress	N/A	N/A	N/A
	Depression	CES-D (8 item)	0-8	Continuous CES-D
	Anxiety	N/A	N/A	N/A
Generation Scotland	Psychological distress	GHQ-28	0-84	Continuous GHQ-28 (Likert scoring)
	Depression	N/A	N/A	N/A
	Anxiety	N/A	N/A	N/A
HBCS	Psychological distress	N/A	N/A	N/A
	Depression	CES-D	0-60	Continuous CES-D

	Anxiety	STAI	20-80	Continuous STAI (trait)
Health 2006	Psychological distress	SCL-90-R	0-4	Continuous sum of depression and anxiety subscales
	Depression	SCL-90-R	0-4	Continuous depression subscale
	Anxiety	SCL-90-R	0-4	Continuous anxiety subscale
Health 2008	Psychological distress	SCL-90-R	0-4	Continuous sum of depression and anxiety subscales
	Depression	SCL-90-R	0-4	Continuous depression subscale
	Anxiety	SCL-90-R	0-4	Continuous anxiety subscale
HUNT	Psychological distress	HADS	0-42	Continuous HADS overall
	Depression	HADS	0-21	Continuous depression subscale
	Anxiety	HADS	0-21	Continuous anxiety subscale
Inter99	Psychological distress	SCL-90-R	0-4	Continuous sum of depression and anxiety subscales
	Depression	SCL-90-R	0-4	Continuous depression subscale
	Anxiety	SCL-90-R	0-4	Continuous anxiety subscale
NFBC 1966	Psychological distress	SCL-25	1-4	Continuous overall SCL-25
	Depression	SCL-25	1-4	Continuous depression subscale
	Anxiety	SCL-25	1-4	Continuous anxiety subscale
NFBC 1986	Psychological distress	YSR	0-40	Continuous average of Withdrawn/depressed subscale and Anxious/depressed subscale
	Depression	YSR	0-14	Continuous YSR Withdrawn/depressed subscale
	Anxiety	YSR	0-26	Continuous YSR Anxious/ depressed subscale
NSHD	Psychological distress	GHQ-28	0-84	Overall GHQ-28 (Likert scoring)

	Depression	N/A	N/A	N/A
	Anxiety	N/A	N/A	N/A
NTR	Psychological distress	ASR	0-42	Continuous average of depression and anxiety subscales
	Depression	ASR	0-28	Continuous ASR Withdrawn/depressed subscale
	Anxiety	ASR	0-14	Continuous ASR Anxious/ depressed subscale
Rotterdam	Psychological distress	N/A	N/A	N/A
	Depression	CES-D	0-60	Continuous CES-D
	Anxiety	N/A	N/A	N/A
SYS-P	Psychological distress	N/A	N/A	N/A
	Depression	CES-D	12-48	Continuous CES-D
	Anxiety	Anxiety questionnaire	1-5	Continuous anxiety scale
Whitehall II	Psychological distress	GHQ-30	0-90	Continuous GHQ-30 (Likert scoring)
	Depression	N/A	N/A	N/A
	Anxiety	N/A	N/A	N/A

CIS-R: Computerised interview schedule- revised, EPDS: Edinburgh Post-natal depression Scale, GHQ: General Health Questionnaire, STAI: State Trait Anxiety Inventory, CES-D: Centre for Epidemiologic Studies Depression, SCL: Symptoms checklist, HADS: Hospital Anxiety and depression scale, YSR: Youth Self Report, ASR: Adult Self Report.

Supplementary Table S3. Descriptive statistics of study populations

					Psychological distress[1]		Psychological distress[2]		Depressed[1]		Depressed[2]		Anxious[1]		Anxious[2]	
	Smoking status	**N**	**% Male**	**Median age (IQR)**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
1958 BC	Never	2,359	50.5	45			242	10.3			164	7.0			152	6.4
	Former	1,451	52.0	45			135	9.3			87	6.0			81	5.6
	Current	1,223	47.8	45			193	15.8			140	11.5			102	8.3
	Total	5,033	50.3	45			570	11.3			391	7.8			335	6.7
ALSPAC Children	Never	1,323	48.9	17.7 (17.5,17.8)	151	11.4	159	12.0	69	5.2	92	7.0	114	8.6	105	7.9
	Current	320	43.1	17.8 (17.6,17.9)	73	22.8	81	25.3	51	15.9	66	20.6	53	16.6	43	13.4
	Total	1,643	47.8	17.7 (17.5,17.8)	224	13.6	240	14.6	120	7.3	158	9.6	167	10.2	148	9.0
ALSPAC Mothers	Never	2,754	0.0	31.8 (28.9,34.8)	379	13.8	294	10.7	258	9.4	197	7.2	309	11.2	230	8.4
	Former	1,330	0.0	31.8 (28.8,34.8)	238	17.9	181	13.6	157	11.8	125	9.4	197	14.8	144	10.8
	Current	1,097	0.0	29.8 (26.8,33)	286	26.1	242	22.1	196	17.9	168	15.3	242	22.1	196	17.9
	Total	5,181	0.0	31.8 (28.8,34.8)	903	17.4	717	13.8	611	11.8	490	9.5	748	14.4	570	11.0
BRHS	Never	984	100.0	64.6 (61.4, 69.3)					76	7.7						
	Former	1,568	100.0	67.0 (62.5, 71.7)					135	8.6						
	Current	410	100.0	65.4 (61.3,69.8)					48	11.7						
	Total	2,962	100.0	66.0 (61.8, 70.8)					259	8.7						
BWHHS	Never	1,812	0.0	69 (64,73)					285	15.7						
	Former	1,015	0.0	69 (64,73)					189	18.6						
	Current	323	0.0	67 (63,72)					76	23.5						
	Total	3,150	0.0	68 (64,73)					550	17.5						
CaPS	Never	248	100.0	55 (52,59)	47	19.0	25	10.1							28	11.3
	Former	484	100.0	57 (53,61)	88	18.2	35	7.2							41	8.5
	Current	381	100.0	56 (53,60)	103	27.0	50	13.1							46	12.1
	Total	1,113	100.0	57 (53,60)	238	21.4	110	9.9							115	10.3

					Psychological distress[1]		Psychological distress[2]		Depressed[1]		Depressed[2]		Anxious[1]		Anxious[2]	
	Smoking status	**N**	**% Male**	**Median age (IQR)**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
CHDS	Never	352	47.7	30	56	15.9			37	10.5			31	8.8		
	Former	164	37.2	30	39	23.8			24	14.6			23	14.0		
	Current	181	53.6	30	53	29.3			31	17.1			40	22.1		
	Total	697	46.8	30	148	21.2			92	13.2			94	13.5		
CoLaus	Never	853	44.6	50 (42.1,57.8)	94	11.0			52	6.1			48	5.6		
	Former	651	53.8	51 (44.3, 59.5)	86	13.2			37	5.7			58	8.9		
	Current	581	52.3	48.3(41.7,56.6)	90	15.5			39	6.7			57	9.8		
	Total	2,085	49.6	49.8 (42.5,58.1)	270	13.0			128	6.1			163	7.8		
ELSA	Never	1,786	33.5	62 (55,70)	244	13.7			186	10.4			82	4.6		
	Former	2,561	54.5	64 (56,72)	381	14.9			317	12.4			95	3.7		
	Current	829	44.8	59 (54,66)	193	23.3			163	19.7			50	6.0		
	Total	5,176	45.7	62 (55,70)	818	15.8			666	12.9			227	4.4		
FINRISK 2002/2007	Never	4,750	35.3	53 (40,62)					288	6.1						
	Former	3,412	57.3	52 (41, 62)					257	7.5						
	Current	2,404	57.3	45 (36,55)					282	11.7						
	Total	10,566	47.4	51 (39,61)					827	7.8						
Generation Scotland	Never	3,990	37.4	56 (49,62)	852	21.4	345	8.7	533	13.4						
	Former	2,493	45.1	59 (52,64)	538	21.6	243	9.8	389	15.6						
	Current	1,002	40.3	52 (45,59)	334	33.3	167	16.7	302	30.1						
	Total	7,485	40.4	57 (49,63)	1,724	23.0	755	10.1	1,224	16.4						
GOYA females	Never	822	0.0	37.7 (35.2,40.8)	61	7.4			47	5.7			20	2.4		
	Former	115	0.0	37.1 (34.1,40.3)	15	13.0			12	10.4			5	4.4		
	Current	177	0.0	36.4 (34.0, 39.5)	23	13.0			22	12.4			4	2.3		
	Total	1,114	0.0	37.5 (34.9, 40.5)	99	8.9			81	7.3			29	2.6		

					Psychological distress[1]		Psychological distress[2]		Depressed[1]		Depressed[2]		Anxious[1]		Anxious[2]	
	Smoking status	**N**	**% Male**	**Median age (IQR)**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
HBCS	Never	642	25.4	61.2 (59.4, 63.7)	134	20.9	100	15.6	110	17.1	69	10.8			61	9.5
	Former	488	55.9	61.1(59.3, 63.7)	92	18.9	58	11.9	72	14.8	34	7.0			41	8.4
	Current	312	46.2	60.2(59.0, 62.2)	82	26.3	57	18.3	73	23.4	39	12.5			36	11.5
	Total	1,442	40.2	60.9 (59.3, 63.4)	308	21.4	215	14.9	255	17.7	142	9.9			138	9.6
Health 2006	Never	1,242	45.3	46.7 (37.5,59.0)			131	10.6			98	7.9			86	6.9
	Former	942	46.5	53.0 (44.1,61.0)			118	12.5			78	8.3			88	9.3
	Current	574	41.1	49.3 (40.7, 58.0)			90	15.7			70	12.2			64	11.2
	Total	2,758	44.8	49.8 (40.3,59.8)			339	12.3			246	8.9			238	8.6
Health 2008	Never	269	47.6	44.0 (38.4,51.8)			26	9.7			21	7.8			15	5.6
	Former	209	40.2	49.8 (42.9, 54.6)			26	12.4			22	10.5			21	10.1
	Current	107	43.0	47.1 (40.4, 53.7)			15	14.0			10	9.4			8	7.5
	Total	585	44.1	46.6 (40.4, 53.3)			67	11.5			53	9.1			44	7.5
HUNT	Never	23,170	42.3	46 (32,62)	4,026	17.4	3,249	14.0	2,064	8.9	2,064	8.9	3,000	13.0	2,027	8.8
	Former	13,619	59.4	52 (42,66)	2,816	20.7	2,334	17.1	1,612	11.8	1,612	11.8	2,007	14.7	1,355	10.0
	Current	16,500	46.3	46 (36,56)	4,043	24.5	3,388	20.5	2,081	12.6	2,081	12.6	3,274	19.8	2,414	14.6
	Total	53,289	47.9	47 (36,61)	10,885	20.4	8,971	16.8	5,757	10.8	5,757	10.8	8,281	15.5	5,796	10.9
Inter 99	Never	576	48.4	44.8 (39.7,50.0)			48	8.3			39	6.8			34	5.9
	Former	388	53.4	49.7 (40.0,54.7)			44	11.3			33	8.5			32	8.3
	Current	529	44.2	44.9 (39.8, 50.0)			100	18.9			75	14.2			68	12.9
	Total	1,493	48.2	45.0 (39.8, 50.0)			192	12.9			147	9.9			134	9.0
NFBC1966	Never	1,891	42.7	31	241	12.7	198	10.5	204	10.8	152	8.0	105	5.6	105	5.6
	Former	609	45.3	31	99	16.3	84	13.8	83	13.6	62	10.2	49	8.1	49	8.1
	Current	1,419	57.3	31	281	19.8	248	17.5	233	16.4	184	13.0	156	11.0	156	11.0
	Total	3,919	48.4	31	621	15.9	530	13.5	520	13.3	398	10.2	310	7.9	310	7.9

					Psychological distress[1]		Psychological distress[2]		Depressed[1]		Depressed[2]		Anxious[1]		Anxious[2]	
	Smoking status	**N**	**% Male**	**Median age (IQR)**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
NFBC1986	Never	733	49.0	16	182	24.8	107	14.6	130	17.7	76	10.4	121	16.5	65	8.9
	Current	410	44.4	16	124	30.2	72	17.6	70	17.1	37	9.0	106	25.9	60	14.6
	Total	1,143	47.3	16	306	26.8	179	15.7	200	17.5	113	9.9	227	19.9	125	10.9
NHANES	Never	357	46.0[3]	28 (22,34)						7.3[3]						
	Former	119	50.0[3]	32 (27,37)						5.9[3]						
	Current	218	53.0[3]	28 (23,34)						9.4[3]						
	Total	694	49.0[3]	29 (23,34)						7.7[3]						
NSHD	Never	1,068	41.7	53	191	17.9	90	8.4								
	Former	893	58.3	53	163	18.3	81	9.1								
	Current	615	50.1	53	132	21.5	71	11.5								
	Total	2,576	49.5	53	486	18.9	242	9.4								
NTR	Never	1,597	31.3	39 (32,55)	162	10.1	219	13.7	112	7.0	143	9.0	107	6.7	162	10.1
	Former	982	38.0	55 (44.3,62)	117	11.9	140	14.3	80	8.2	96	9.8	78	7.9	100	10.2
	Current	469	37.1	48 (35,58)	79	16.8	85	18.1	57	12.2	63	13.4	56	11.9	65	13.9
	Total	3,048	34.4	47 (35,59)	358	11.8	444	14.6	249	8.2	302	9.9	241	7.9	327	10.7
Patch 2	Former	100	46.0	53(46,65.5)					36	36.0						
	Current	132	34.1	54 (47,61)					51	38.7						
	Total	232	39.2	53.5(47,62)					87	37.5						
Rotterdam	Never	2,845	30.2	61.8 (56.2,73.2)	423	14.9			275	9.7			233	8.2		
	Former	2,531	57.4	72.2 (66.7,77.8)	364	14.4			227	9.0			201	7.9		
	Current	1,144	45.8	62.0 (56.3,70.5)	230	20.1			172	15.0			116	10.1		
	Total	6,520	43.5	67.5 (59.9, 75.5)	1017	15.6			674	10.3			550	8.4		
SYS-P	Never	250	51.2	43(40,46)	33	13.2	32	12.8	13	5.2	15	6.0	23	9.2	23	9.2
	Former	374	48.1	44(41,47)	42	11.2	52	13.9	18	4.8	30	8.0	31	8.3	34	9.1
	Current	260	46.9	42 (39, 45.3)	45	17.3	50	19.2	25	9.6	33	12.7	29	11.2	29	11.2
	Total	884	48.6	43(40,46)	120	13.6	134	15.2	56	6.3	78	8.8	83	9.4	86	9.4

					Psychological distress[1]		Psychological distress[2]		Depressed[1]		Depressed[2]		Anxious[1]		Anxious[2]	
	Smoking status	**N**	**% Male**	**Median age (IQR)**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Whitehall II	Never	1,503	73.1	42 (38,49)	403	26.8	148	9.9								
	Former	930	78.9	44 (39,49)	241	25.9	96	10.3								
	Current	411	69.8	42 (38,48)	124	30.2	49	11.9								
	Total	2,844	74.5	43 (39,49)	768	27.0	293	10.3								

1. Case definition 1
2. Case definition 2.
3. NHANES data are weighted percentages

Supplementary Table S4. Allele frequencies for rs1051730/rs16969968 in the CARTA studies

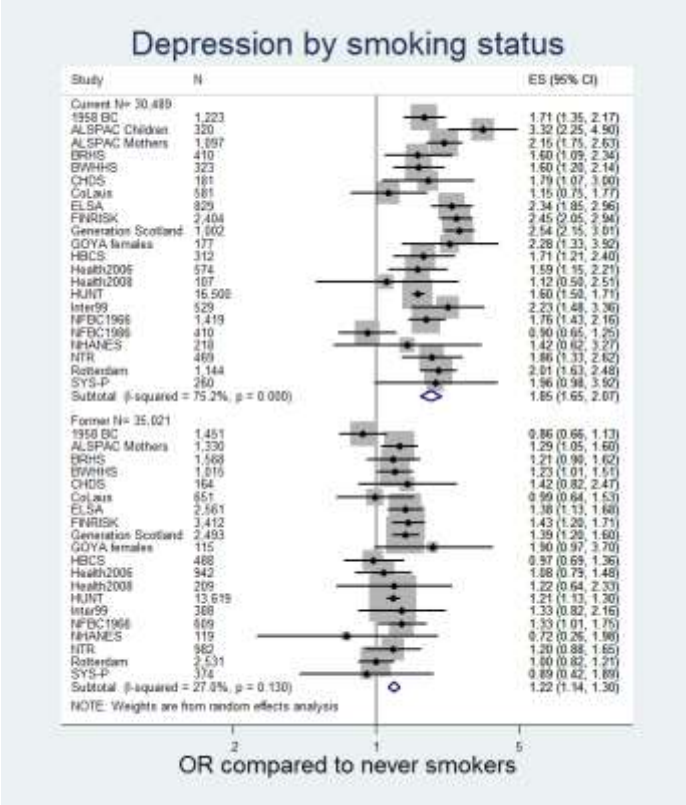
Study	SNP	Total	Major homozygotes	Heterozygotes	Minor homozygotes	MAF	HWE p-value[1]
1958 BC	rs16969968	5,033	2,184	2,233	616	0.34	0.22
ALSPAC Children	rs1051730	1,643	732	738	173	0.33	0.54
ALSPAC Mothers	rs1051730	5,181	2,290	2,295	596	0.34	0.58
BRHS	rs1051730	2,962	1,350	1,280	332	0.33	0.28
BWHHS	rs1051730	3,150	1,385	1,419	346	0.34	0.55
CaPS	rs16969968	1,113	515	487	111	0.32	0.84
CHDS	rs16969968	697	320	312	65	0.32	0.43
CoLaus	rs1051730	2,085	879	942	264	0.35	0.63
ELSA	rs16969968	5,176	2,355	2,261	560	0.33	0.61
FINRISK 2002/2007	rs16969968	10,566	4,834	4,625	1,107	0.32	0.98
Generation Scotland	rs1051730	7,485	3,346	3,339	800	0.33	0.45
GOYA females	rs1051730	1,114	476	526	112	0.34	0.06
HBCS	rs1051730	1,442	635	645	162	0.34	0.95
Health 2006	rs1051730	2,758	1,248	1,206	304	0.33	0.64
Health 2008	rs16969968	585	271	255	59	0.32	1.00
HUNT	rs1051730	53,289	23,647	23,591	6,051	0.33	0.15
Inter 99	rs1051730	1,493	660	667	166	0.33	0.95
NFBC1966	rs1051730	3,919	1,809	1,692	418	0.32	0.47
NFBC1986	rs1051730	1,143	540	508	95	0.31	0.12
NHANES	rs1051730	694	0.38[2]	0.48[2]	0.14[2]	0.38[2]	-
NSHD	rs16969968	2,576	1,189	1,136	251	0.32	0.41
NTR	rs16969968	3,048	1,427	1,310	311	0.32	0.68
Patch 2	rs1051730	232	101	95	36	0.36	0.09
Rotterdam	rs1051730	6,520	3,044	2,821	655	0.32	0.98
SYS-P	rs1051730	884	329	424	131	0.39	0.83
Whitehall II	rs16969968	2,844	1,269	1,270	305	0.33	0.64

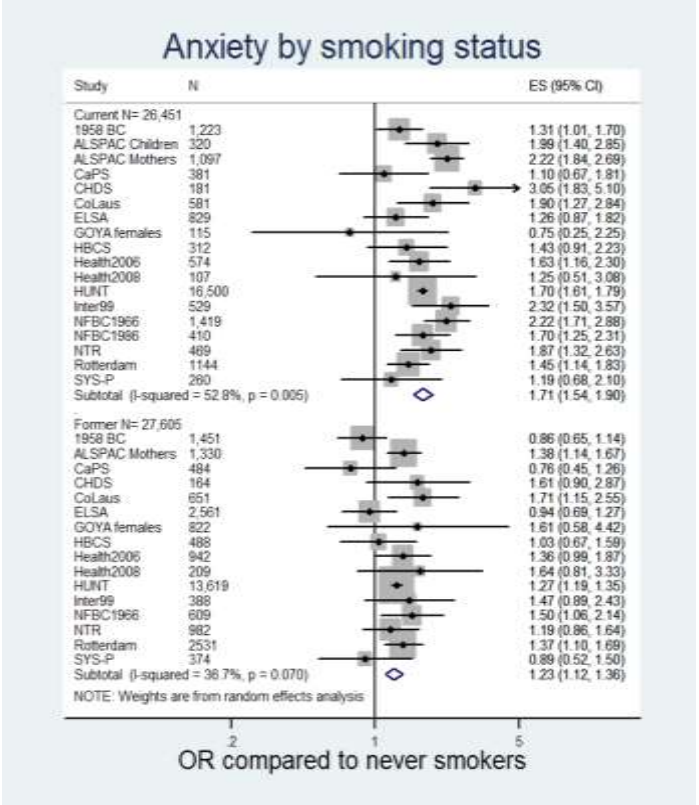
MAF: Minor allele frequency, HWE: Hardy Weinberg Equilibrium

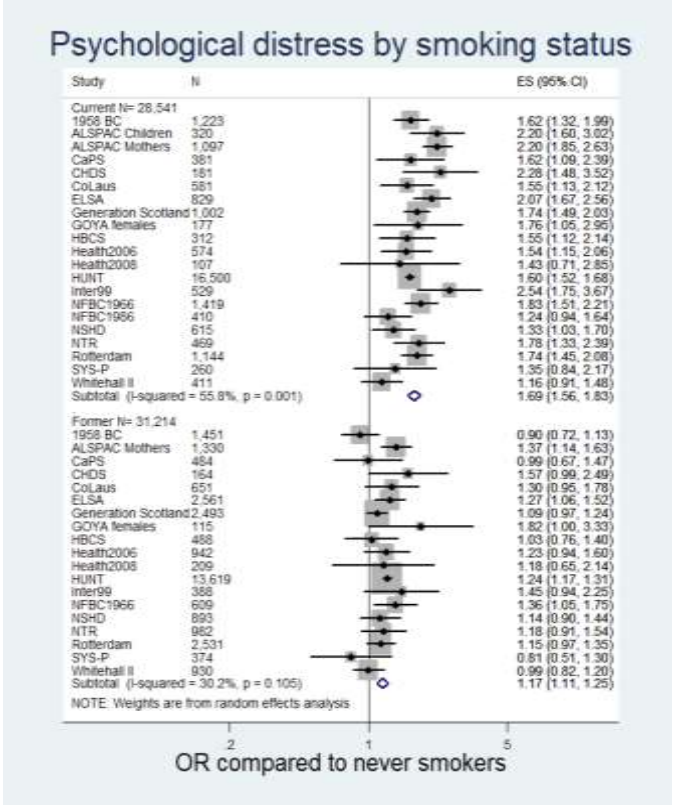
1. P-value from chi-square exact test

2. Weighted percentages of genotypes are presented from NHANES as this is survey data

Supplementary Figure S1. Age- and sex- adjusted association of smoking status with depression, anxiety and psychological distress (random effects meta-analysis)

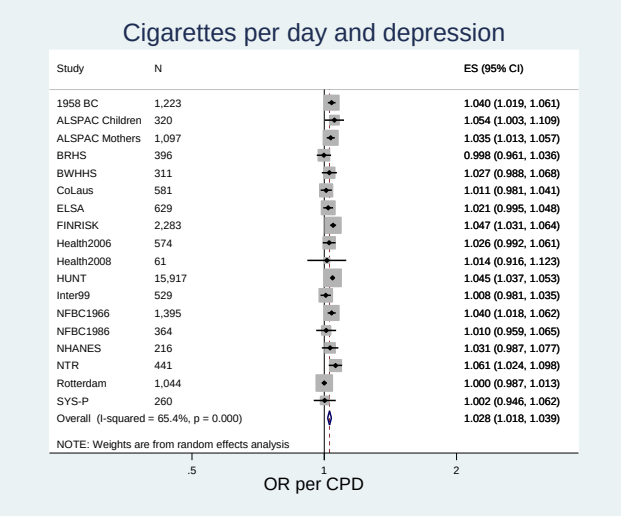


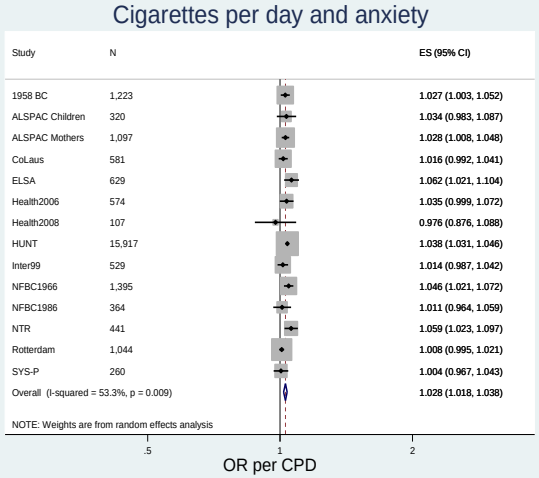


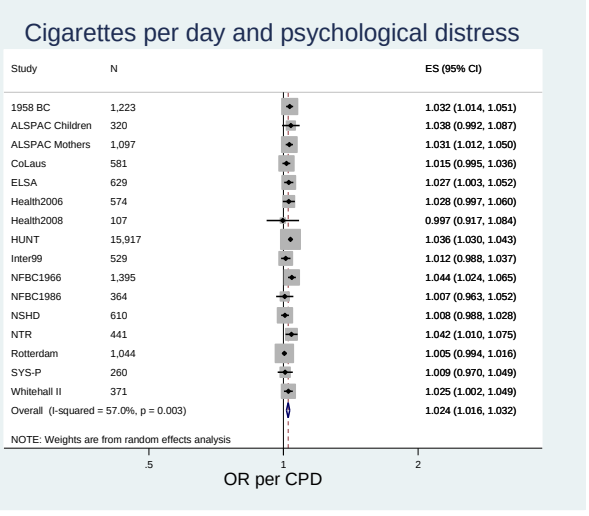


N in figure represents number of current, former and ever smokers. N for never smokers (reference group) by study is as follows: 1958 BC: 2,359, ALSPAC children: 1,323, ALSPAC mothers: 2,754, BRHS: 984, BWHHS: 1,812, CaPS: 248, CHDS: 352, CoLaus: 853, ELSA: 1,786, FINRISK: 4,750, Generation Scotland: 3,990, GOYA females: 822, HBCS: 642, Health2006: 1,242, Health2008: 269, HUNT:23,170, Inter99: 576, NFBC1966: 1,891, NFBC1986: 733, NHANES: 357, NSHD: 1,068, NTR: 1,597, Rotterdam: 2,845, SYS-P: 250, Whitehall II: 1,503.

Supplementary Figure S2. Age- and sex- adjusted association of smoking heaviness with depression, anxiety and psychological distress (random effects meta-analysis)

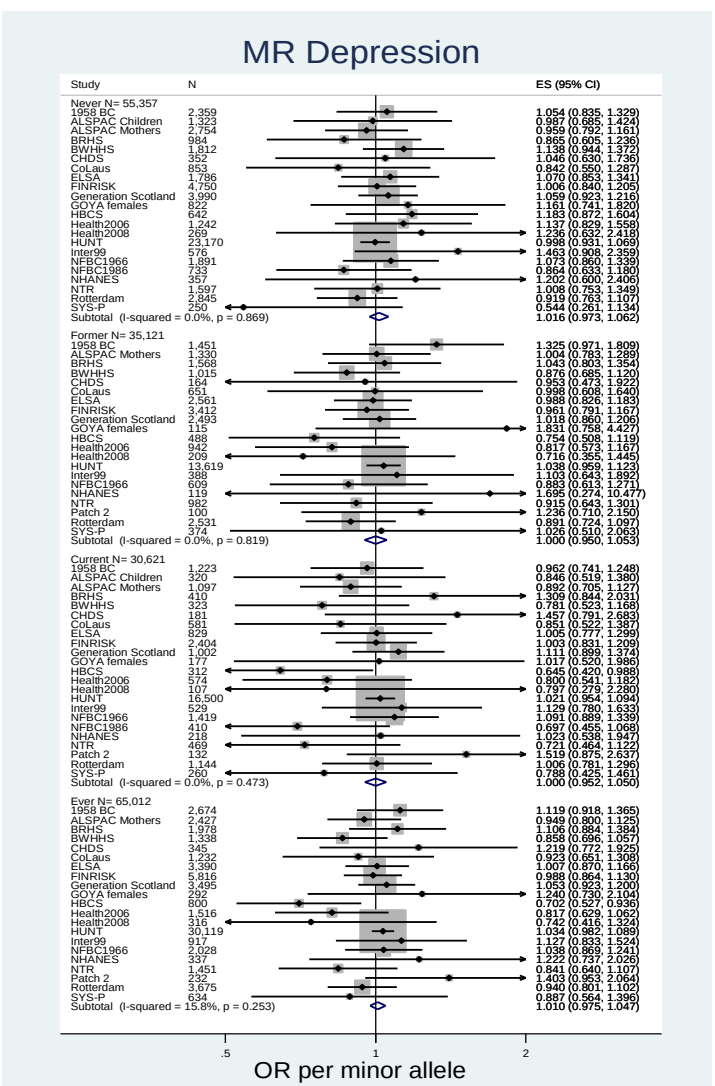


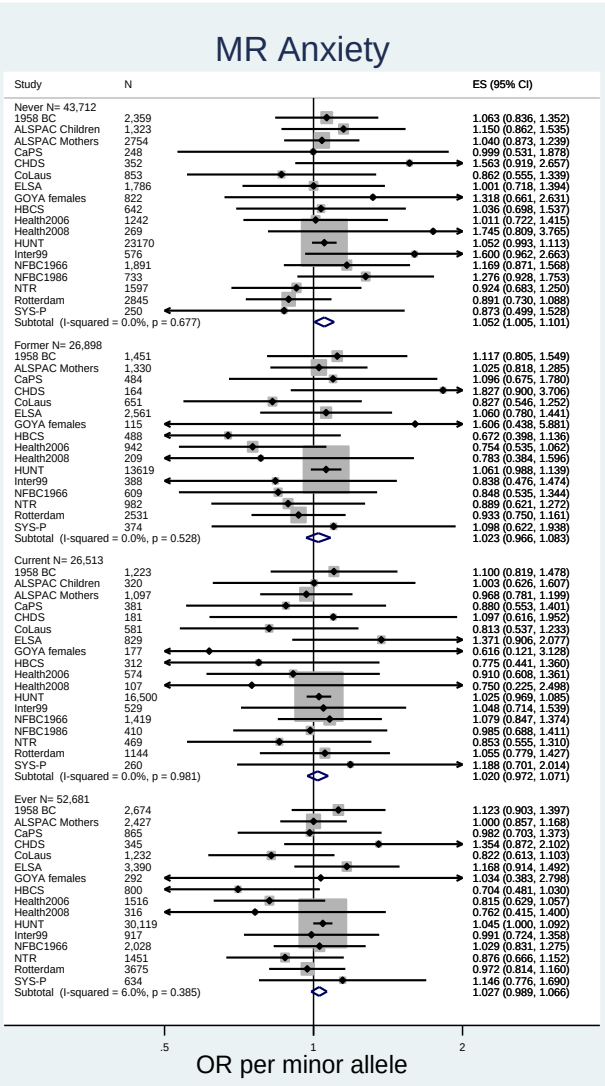


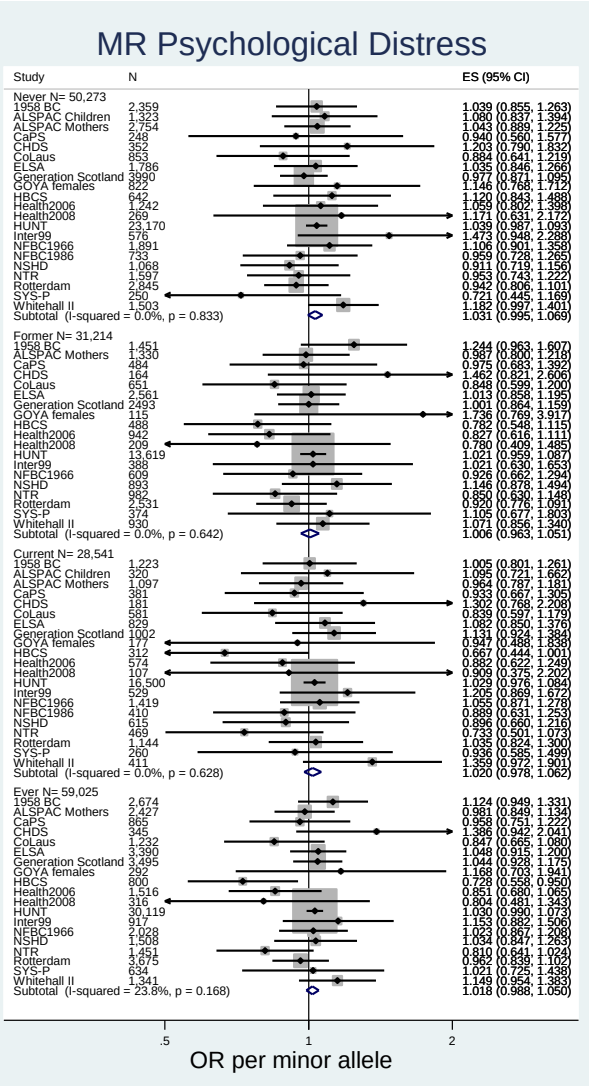


CPD: cigarettes per day. N for depression analysis= 27,641 N for anxiety analysis=24,481, N for psychological distress analysis= 25,462

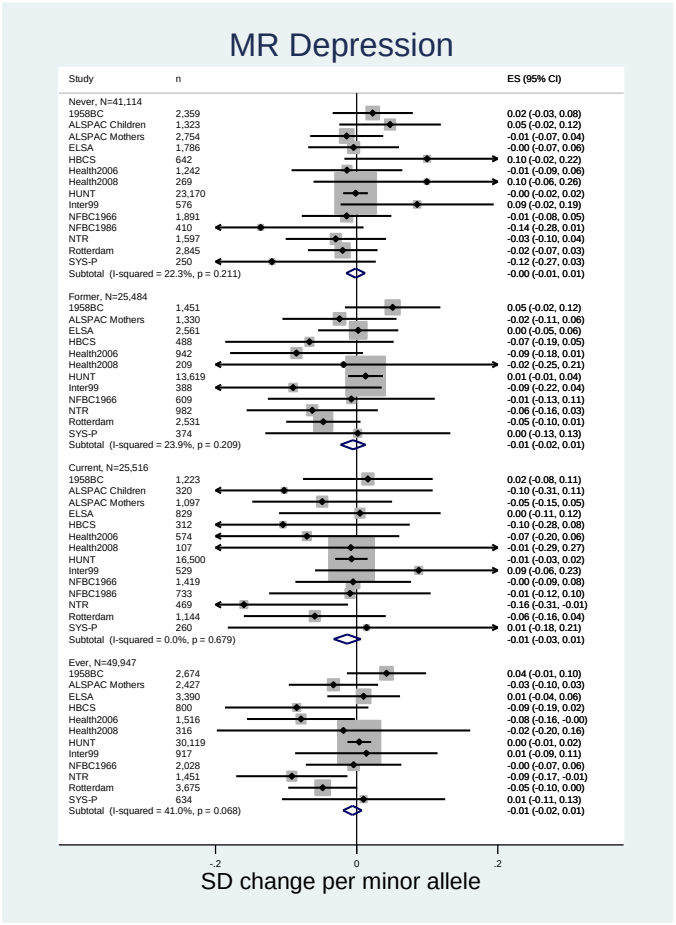
Supplementary Figure S3 **Age- and sex- adjusted association of rs16969968/rs1051730 with depression, anxiety and psychological distress (fixed effects meta-analysis)**

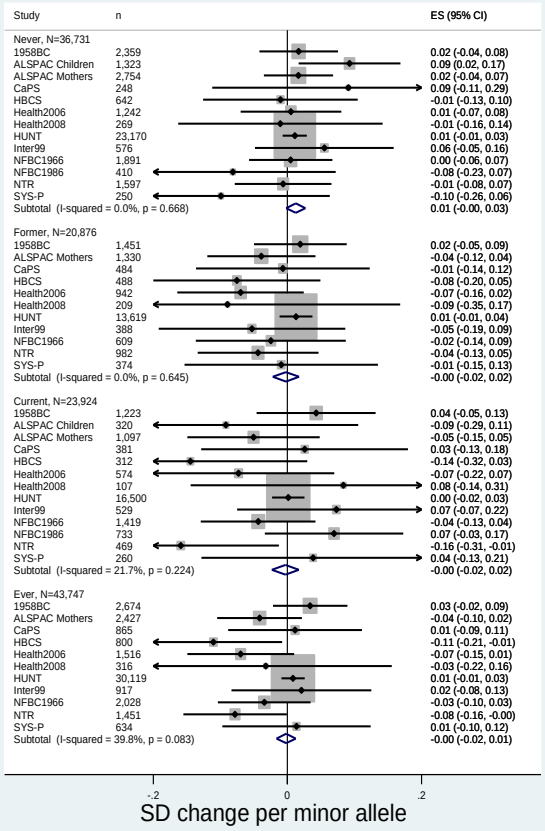


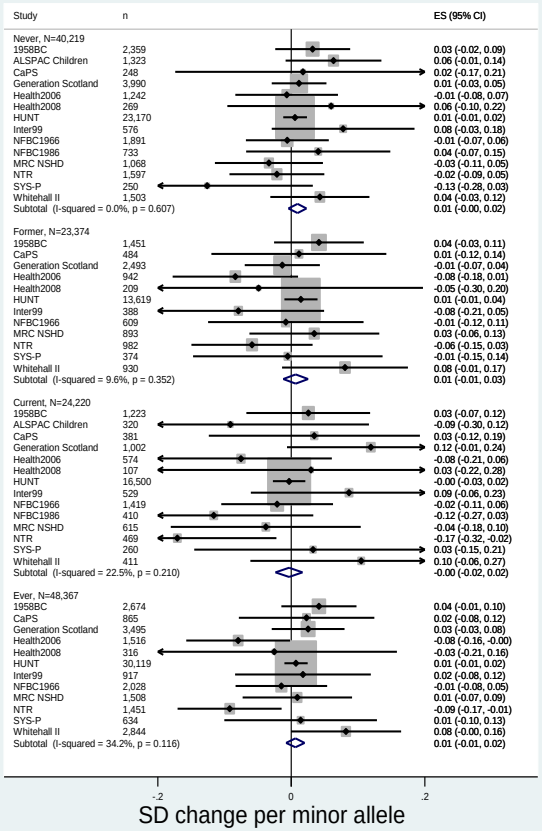




Supplementary Figure S4. Association of rs1051730/rs16969968 with continuous scales of depression, anxiety and psychological distress

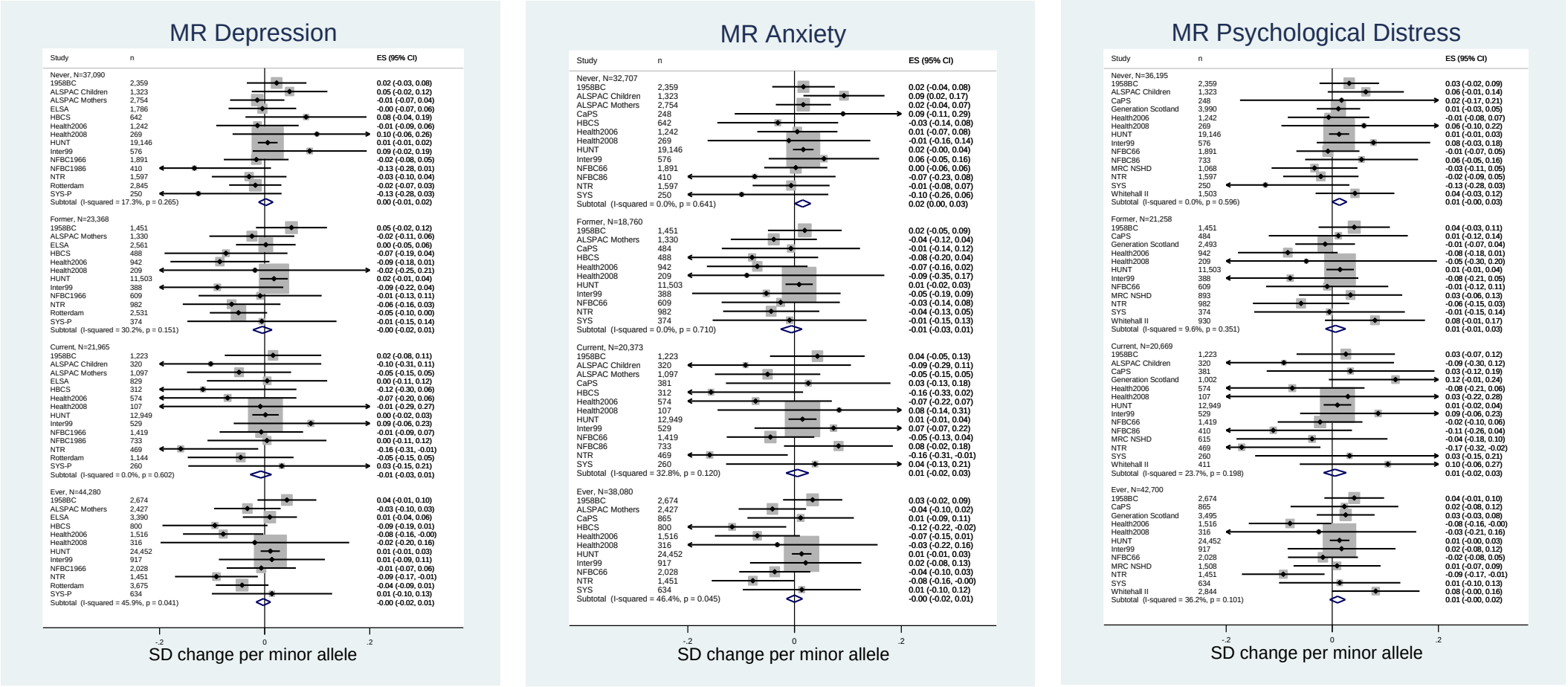






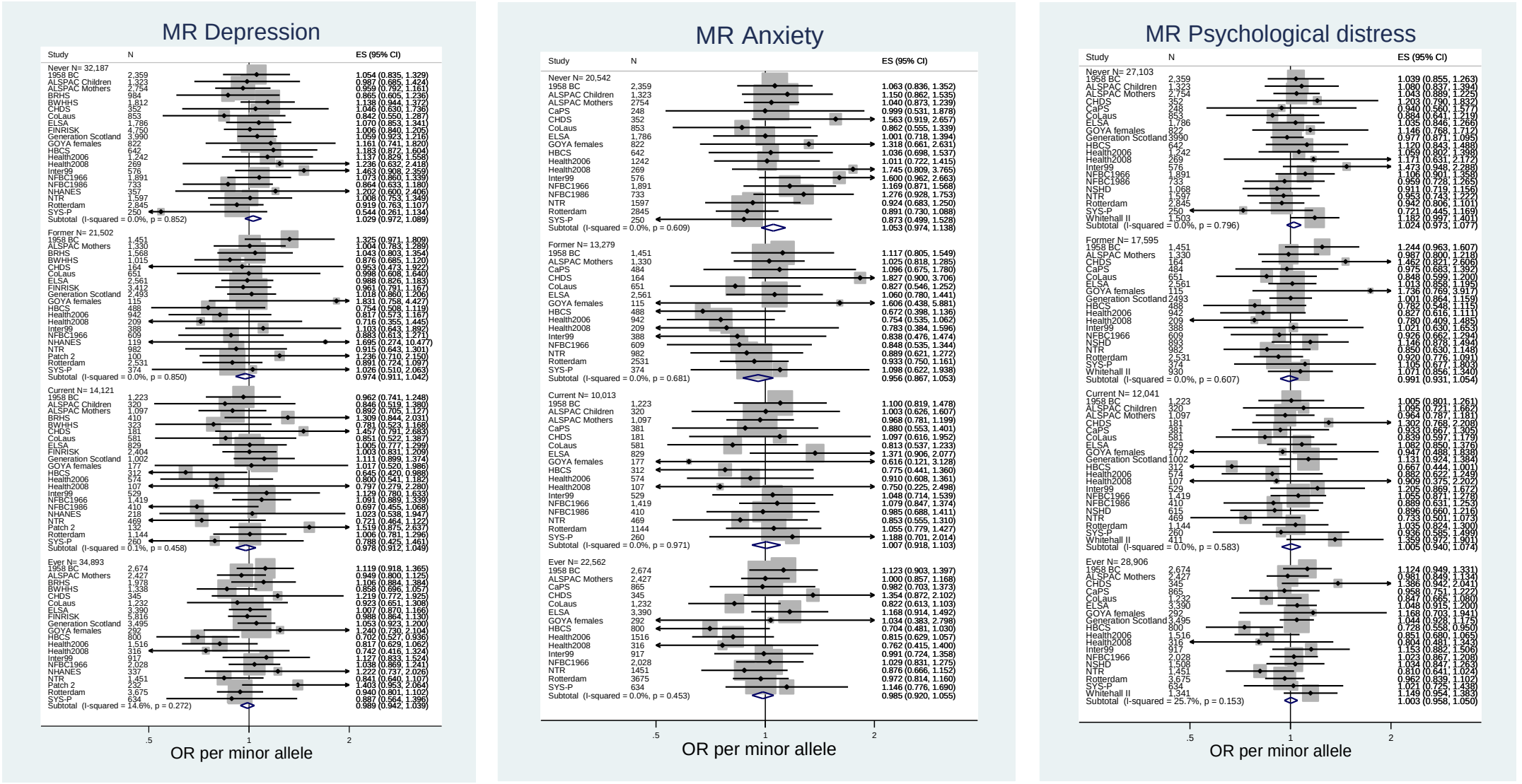
SD: standard deviation. Associations from linear regression, adjusted for age and sex. All symptom scales converted to z-scores.

Supplementary Figure S5 **Medication adjusted association of rs1051730/rs16969968 with continuous scales of depression, anxiety and psychological distress**

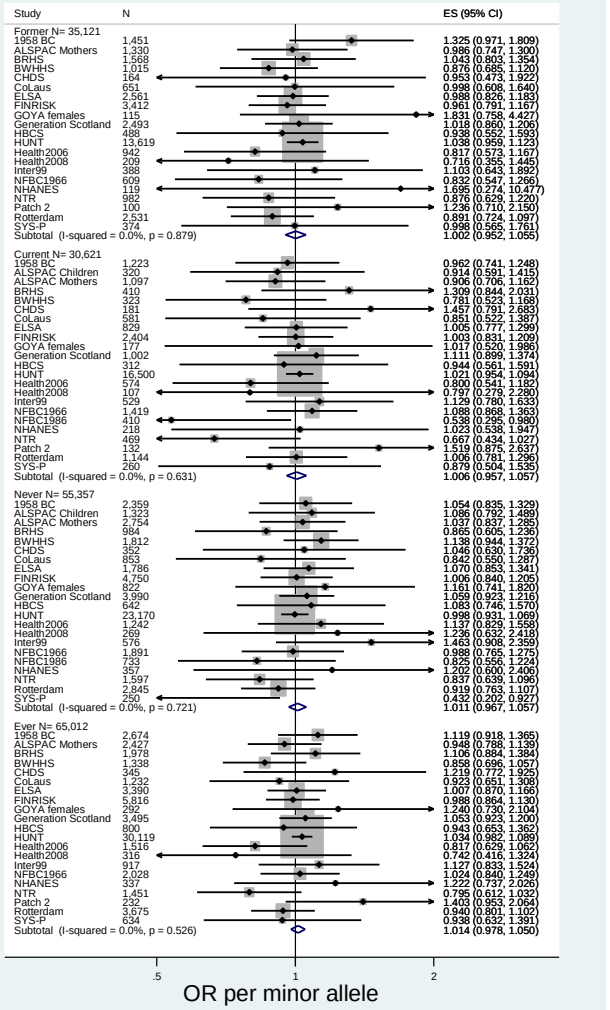


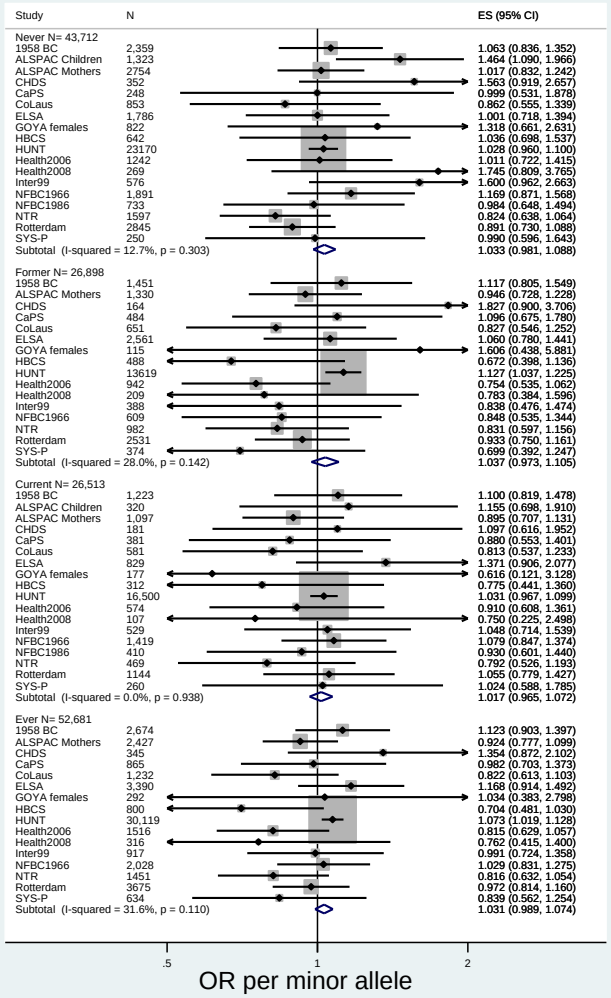
SD: standard deviation. Associations from linear regression, adjusted for age, sex and use of antidepressant medication in HUNT, NFBC1966, NFBC1986 and Rotterdam. All symptom scales converted to z-scores.

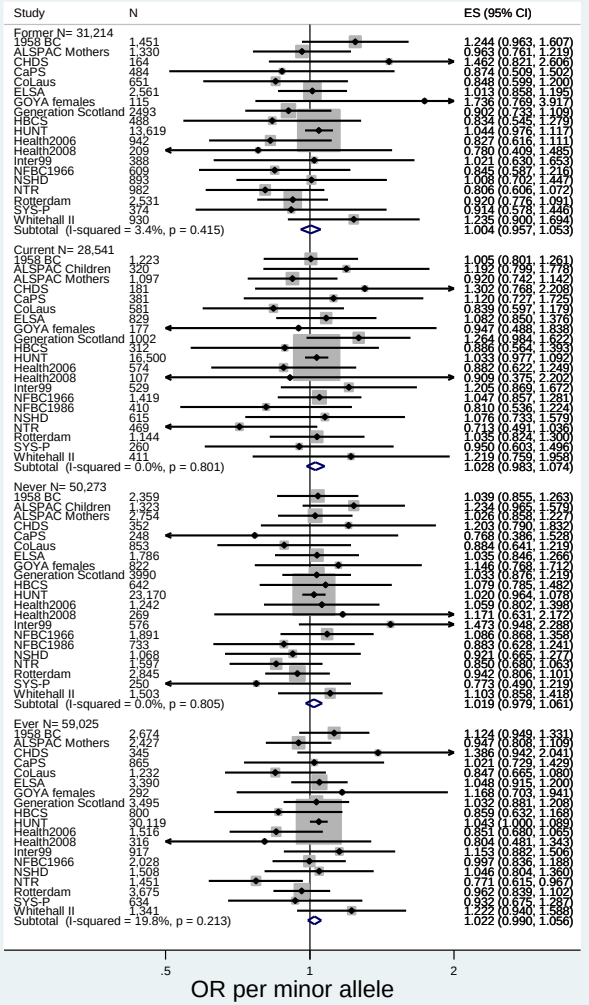
Supplementary Figure S6. **Age- and sex- adjusted association of rs16969968/rs1051730 with anxiety excluding HUNT (fixed effects meta-analysis)**



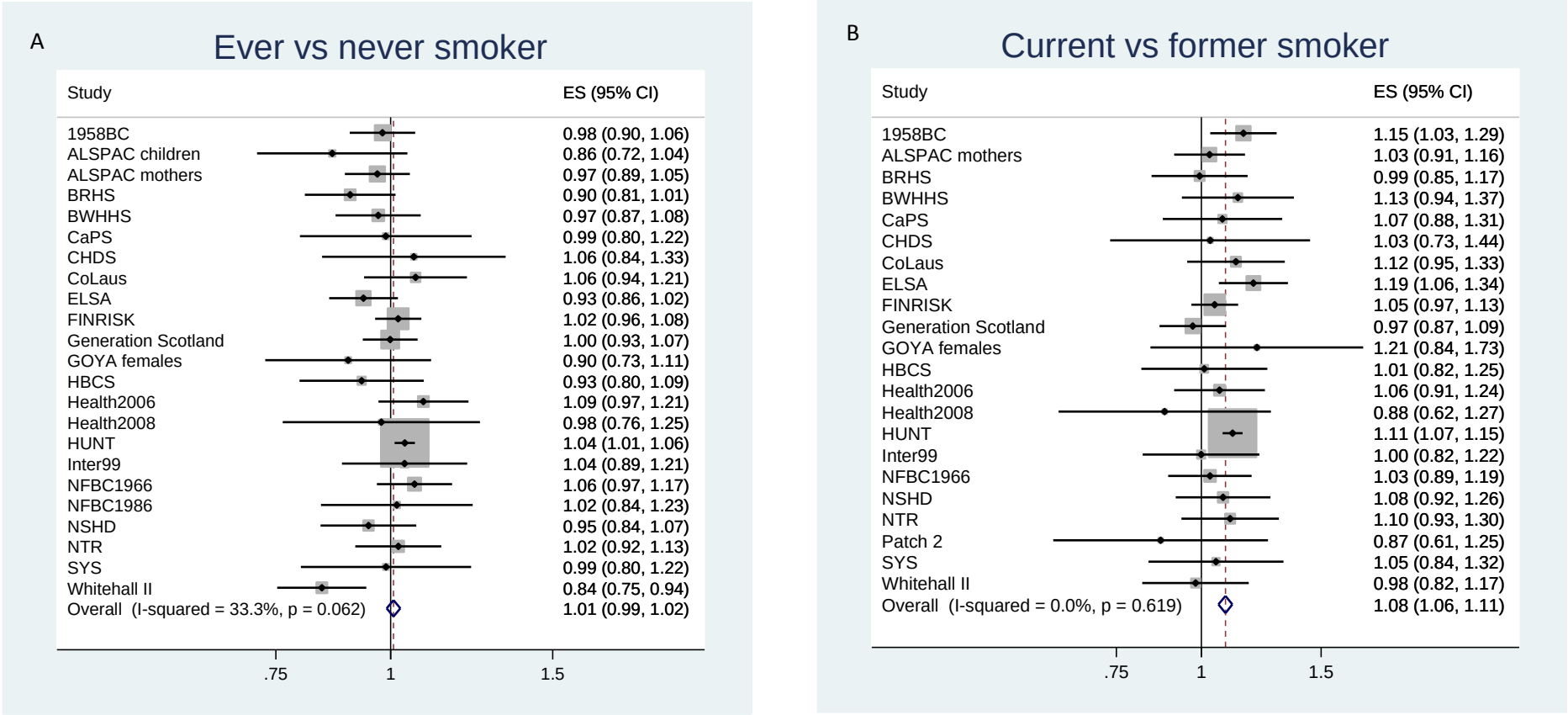
Supplementary Figure S7. **Age- and sex- adjusted association of rs16969968/rs1051730 with anxiety using case definition 2 in preference (fixed effects meta-analysis)**





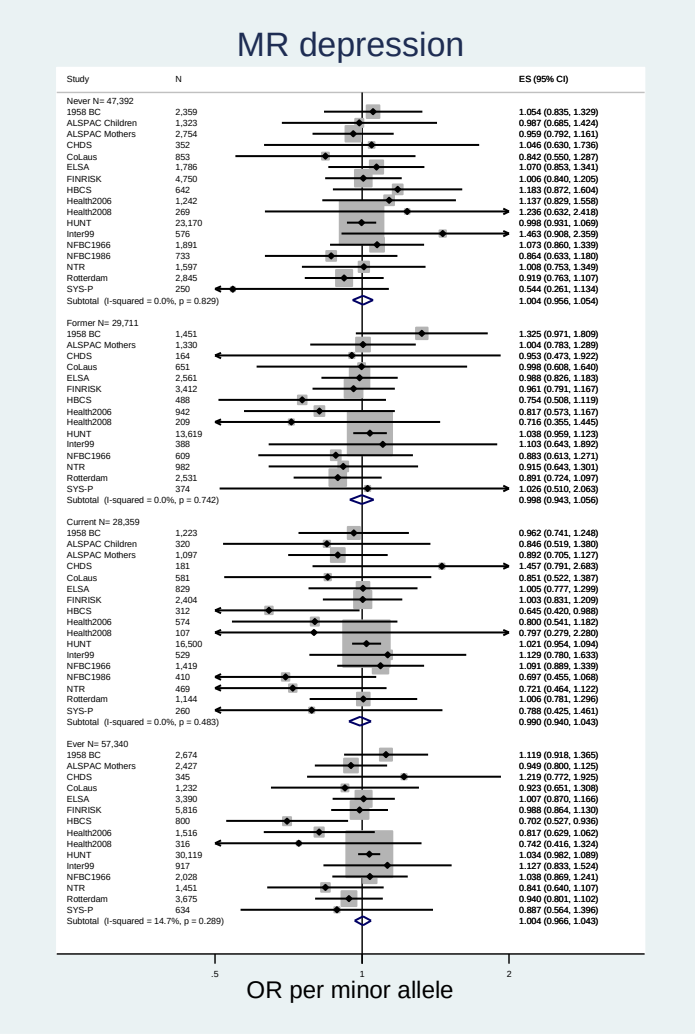


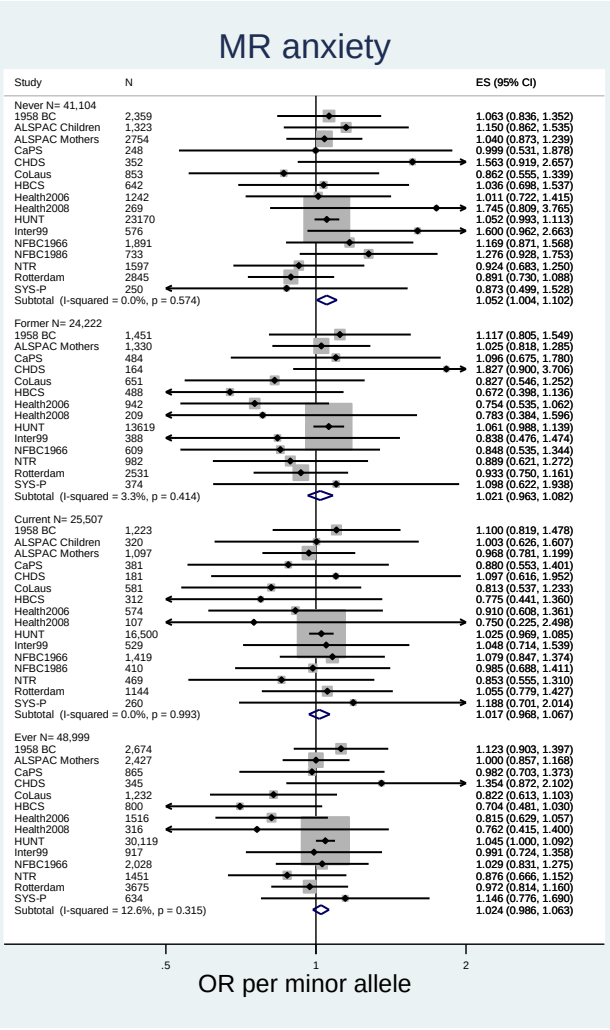
Supplementary Figure S8. Associations of smoking status with rs1051730/rs16969968 in CARTA



Effect sizes presented as odds ratios. Never smokers reference group in A. Former smokers reference group in B. Includes all samples apart from Rotterdam and NHANES as genotype by smoking status was not available for these samples

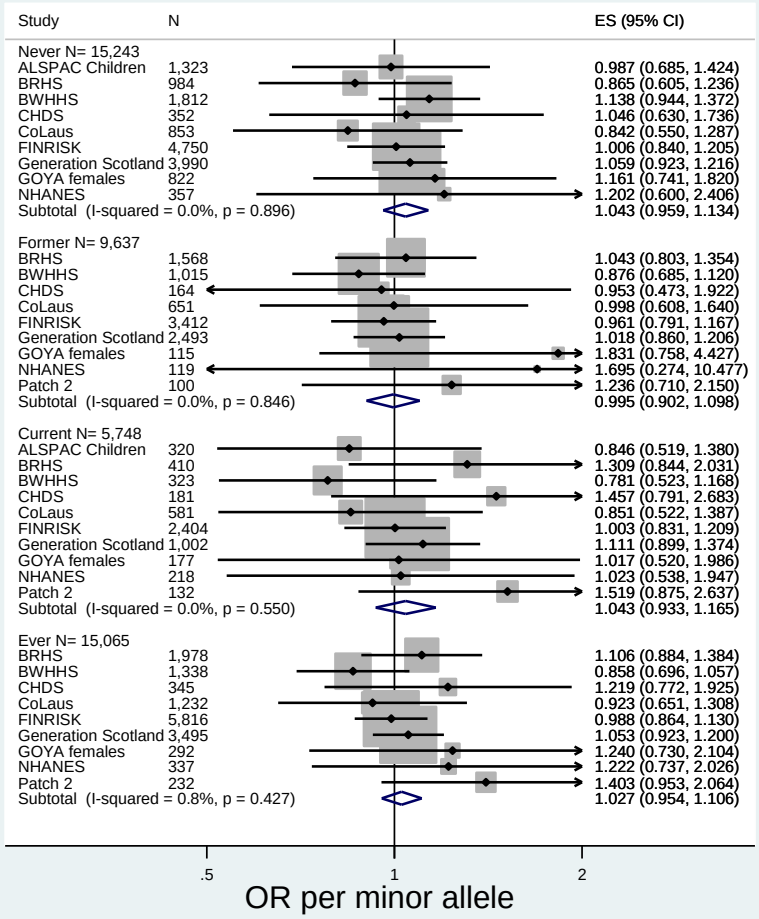
Supplementary Figure S9. **Age- and sex- adjusted association of rs16969968/rs1051730 with depression and anxiety excluding lifetime diagnosis (fixed effects meta-analysis)**

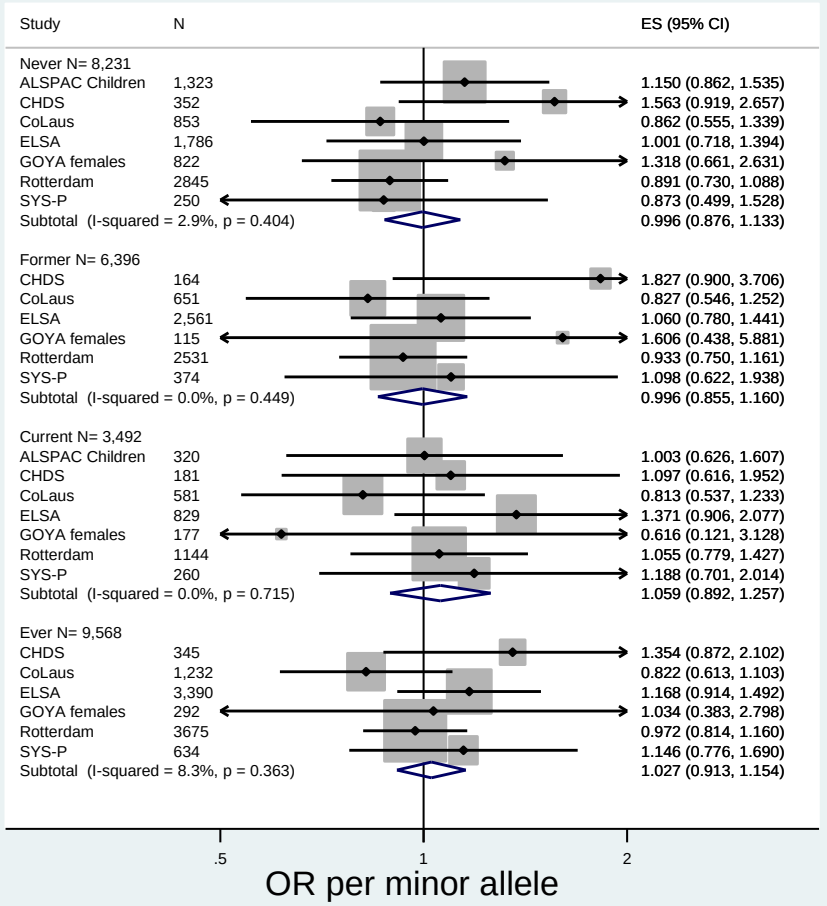




Depression analysis excludes BRHS, BWHHS, Generation Scotland, GOYA females, NHANES, Patch 2. Anxiety analysis excludes ELSA and GOYA females.

Supplementary Figure S10. **Age- and sex- adjusted association of rs16969968/rs1051730 with depression and anxiety restricting to questionnaires using clinical criteria or self-reported diagnosis (fixed effects meta-analysis)**





Depression analysis excludes 1958BC, ALSPAC Mothers, ELSA,HBCS, Health2006, Health2008, HUNT, Inter99, NFBC1966, NFBC1966, NTR, Rotterdam, SYS-P.
Anxiety analysis excludes 1958BC, ALSPAC Mothers, CaPS, HBCS, Health2006, Health2008, HUNT, Inter99, NFBC1966, NFBC1966, NTR.

1. Power C, Elliott J. Cohort profile: 1958 British birth cohort (National Child Development Study). International journal of epidemiology 2006;35(1):34-41 doi: 10.1093/ije/dyi183[published Online First: Epub Date]|.
2. Atherton K, Fuller E, Shepherd P, et al. Loss and representativeness in a biomedical survey at age 45 years: 1958 British birth cohort. Journal of epidemiology and community health 2008;62(3):216-23 doi: 10.1136/jech.2006.058966[published Online First: Epub Date]|.
3. International Multiple Sclerosis Genetics C, Wellcome Trust Case Control C, Sawcer S, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 2011;476(7359):214-9 doi: 10.1038/nature10251[published Online First: Epub Date]|.
4. Barrett JC, Clayton DG, Concannon P, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nature genetics 2009;41(6):703-7 doi: 10.1038/ng.381[published Online First: Epub Date]|.
5. Lewis G, Pelosi AJ, Araya R, et al. Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. Psychological medicine 1992;22(2):465-86
6. Boyd A, Golding J, Macleod J, et al. Cohort Profile: The 'Children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. International journal of epidemiology 2013;42(1):111-27 doi: 10.1093/ije/dys064[published Online First: Epub Date]|.
7. Freathy RM, Ring SM, Shields B, et al. A common genetic variant in the 15q24 nicotinic acetylcholine receptor gene cluster (CHRNA5-CHRNA3-CHRNB4) is associated with a reduced ability of women to quit smoking in pregnancy. Human molecular genetics 2009;18(15):2922-27 doi: Doi 10.1093/Hmg/Ddp216[published Online First: Epub Date]|.
8. Lewis SJ, Araya R, Smith GD, et al. Smoking is associated with, but does not cause, depressed mood in pregnancy--a mendelian randomization study. PloS one 2011;6(7):e21689 doi: 10.1371/journal.pone.0021689[published Online First: Epub Date]|.
9. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. The British journal of psychiatry : the journal of mental science 1987;150:782-6
10. Crown S, Crisp AH. A short clinical diagnostic self-rating scale for psychoneurotic patients. The Middlesex Hospital Questionnaire (M.H.Q.). The British journal of psychiatry : the journal of mental science 1966;112(490):917-23
11. Heron J, O'Connor TG, Evans J, et al. The course of anxiety and depression through pregnancy and the postpartum in a community sample. Journal of affective disorders 2004;80(1):65-73 doi: 10.1016/j.jad.2003.08.004[published Online First: Epub Date]|.
12. Jones RW, Ring S, Tyfield L, et al. A new human genetic resource: a DNA bank established as part of the Avon longitudinal study of pregnancy and childhood (ALSPAC). European journal of human genetics : EJHG 2000;8(9):653-60 doi: 10.1038/sj.ejhg.5200502[published Online First: Epub Date]|.
13. Lewis G. Computerized assessments of psychiatric disorder using PROQSY: discussion paper. Journal of the Royal Society of Medicine 1992;85(7):403-6
14. Walker M, Whincup PH, Shaper AG. The British Regional Heart Study 1975-2004. International journal of epidemiology 2004;33(6):1185-92 doi: 10.1093/ije/dyh295[published Online First: Epub Date]|.
15. Lawlor DA, Bedford C, Taylor M, et al. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. J Epidemiol Community Health 2003;57(2):134-40
16. Caerphilly and Speedwell collaborative heart disease studies. The Caerphilly and Speedwell Collaborative Group. Journal of epidemiology and community health 1984;38(3):259-62
17. Goldberg D, Williams P. *A User's Guide to the General Health Questionnaire*. Windsor: NFER-NELSON, 2006.

18. Williams MD, Harris R, Dayan CM, et al. Thyroid function and the natural history of depression: findings from the Caerphilly Prospective Study (CaPS) and a meta-analysis. Clinical endocrinology 2009;70(3):484-92 doi: 10.1111/j.1365-2265.2008.03352.x[published Online First: Epub Date]|.
19. Head J, Stansfeld SA, Ebmeier KP, et al. Use of self-administered instruments to assess psychiatric disorders in older people: validity of the General Health Questionnaire, the Center for Epidemiologic Studies Depression Scale and the self-completion version of the revised Clinical Interview Schedule. Psychological medicine 2013;43(12):2649-56 doi: 10.1017/S0033291713000342[published Online First: Epub Date]|.
20. Spielberger CD. The State-Trait Anxiety Inventory. In: McDowell I, ed. Measuring Health: A Guide to Rating Scales and Questionnaires: Oxford University Press, 2006.
21. Fergusson DM, Horwood LJ. The Christchurch Health and Development Study: Review of findings on child and adolescent mental health. Aust. N.Z. J. Psychiatry 2001;35(3):287-96
22. Fergusson DM, Horwood LJ, Shannon FT, et al. The Christchurch Child Development Study: A review of epidemiological findings. Paediatr. Perinat. Epidemiol. 1989;3:302-25
23. Ciulla TA, Sklar RM, Hauser SL. A simple method for DNA purification from peripheral blood. Analytical biochemistry 1988;174(2):485-8
24. WHO. Composite International Diagnostic Interview (CIDI). 1993.
25. Firmann M, Mayor V, Vidal PM, et al. The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. BMC cardiovascular disorders 2008;8:6 doi: 10.1186/1471-2261-8-6[published Online First: Epub Date]|.
26. Preisig M, Waeber G, Vollenweider P, et al. The PsyCoLaus study: methodology and characteristics of the sample of a population-based survey on psychiatric disorders and their association with genetic and cardiovascular risk factors. BMC psychiatry 2009;9:9 doi: 10.1186/1471-244X-9-9[published Online First: Epub Date]|.
27. Nurnberger JI, Jr., Blehar MC, Kaufmann CA, et al. Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. Archives of general psychiatry 1994;51(11):849-59; discussion 63-4
28. Preisig M, Fenton BT, Matthey ML, et al. Diagnostic interview for genetic studies (DIGS): inter-rater and test-retest reliability of the French version. European archives of psychiatry and clinical neuroscience 1999;249(4):174-9
29. Steptoe A, Breeze E, Banks J, et al. Cohort Profile: The English Longitudinal Study of Ageing. International journal of epidemiology 2012 doi: 10.1093/ije/dys168[published Online First: Epub Date]|.
30. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. . Applied Psychological Measurement 1977;1:385-401
31. Steffick DE. Documentation of Affective Functioning Measures in the Health and Retirement Study. HRS Health Working Group, Ann Arbor, MI, DR-005, 2000.
32. Vartiainen E, Laatikainen T, Peltonen M, et al. Thirty-five-year trends in cardiovascular risk factors in Finland. International journal of epidemiology 2010;39(2):504-18 doi: 10.1093/ije/dyp330[published Online First: Epub Date]|.
33. Hallfors J, Loukola A, Pitkaniemi J, et al. Scrutiny of the CHRNA5-CHRNA3-CHRNB4 smoking behavior locus reveals a novel association with alcohol use in a Finnish population based study. International journal of molecular epidemiology and genetics 2013;4(2):109-19
34. Smith BH, Campbell A, Linksted P, et al. Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. International journal of epidemiology 2013;42(3):689-700 doi: 10.1093/ije/dys084[published Online First: Epub Date]|.
35. Kerr SM, Campbell A, Murphy L, et al. Pedigree and genotyping quality analyses of over 10,000 DNA samples from the Generation Scotland: Scottish Family Health Study. BMC medical genetics 2013;14:38 doi: 10.1186/1471-2350-14-38[published Online First: Epub Date]|.
36. Gunderson KL. Whole-genome genotyping on bead arrays. Methods in molecular biology 2009;529:197-213 doi: 10.1007/978-1-59745-538-1_13[published Online First: Epub Date]|.

37. First MB, Gibbon M, Spitzer RL, et al. *Structured clinical interview for DSM-IV-TR Axis Disorders (Research Version).* New York: Biometrics Research: New York State Psychiatric Institute, 2002.
38. Mishra GD, McNaughton SA, O'Connell MA, et al. Intake of B vitamins in childhood and adult life in relation to psychological distress among women in a British birth cohort. Public health nutrition 2009;12(2):166-74 doi: 10.1017/S1368980008002413[published Online First: Epub Date]|.
39. Nohr EA, Timpson NJ, Andersen CS, et al. Severe obesity in young women and reproductive health: the Danish National Birth Cohort. PloS one 2009;4(12):e8444 doi: 10.1371/journal.pone.0008444[published Online First: Epub Date]|.
40. Paternoster L, Evans DM, Nohr EA, et al. Genome-wide population-based association study of extremely overweight young adults--the GOYA study. PloS one 2011;6(9):e24303 doi: 10.1371/journal.pone.0024303[published Online First: Epub Date]|.
41. Barker DJ, Osmond C, Forsen TJ, et al. Trajectories of growth among children who have coronary events as adults. The New England journal of medicine 2005;353(17):1802-9 doi: 10.1056/NEJMoa044160[published Online First: Epub Date]|.
42. Eriksson JG, Osmond C, Kajantie E, et al. Patterns of growth among children who later develop type 2 diabetes or its risk factors. Diabetologia 2006;49(12):2853-8 doi: 10.1007/s00125-006-0459-1[published Online First: Epub Date]|.
43. Rikkonen K, Pesonen AK, Heinonen K, et al. Infant growth and hostility in adult life. Psychosomatic medicine 2008;70(3):306-13 doi: 10.1097/PSY.0b013e3181651638[published Online First: Epub Date]|.
44. Radloff LS. The use of the Center for Epidemiologic Studies Depression Scale in adolescents and young adults. Journal of youth and adolescence 1991;20(2):149-66 doi: 10.1007/BF01537606[published Online First: Epub Date]|.
45. Beekman AT, Deeg DJ, Van Limbeek J, et al. Criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D): results from a community-based sample of older subjects in The Netherlands. Psychological medicine 1997;27(1):231-5
46. Thuesen BH, Cerqueira C, Aadahl M, et al. Cohort Profile: The Health2006 cohort, Research Centre for Prevention and Health. International journal of epidemiology 2013 doi: 10.1093/ije/dyt009[published Online First: Epub Date]|.
47. Derogatis LR. *SCL-90-R: Symptom Checklist-90-R. Administration, Scoring and Procedures Manual*. 1st ed. Minneapolis: National Computer Systems, 1977.
48. Byberg S, Hansen AL, Christensen DL, et al. Sleep duration and sleep quality are associated differently with alterations of glucose homeostasis. Diabetic medicine : a journal of the British Diabetic Association 2012;29(9):e354-60 doi: 10.1111/j.1464-5491.2012.03711.x[published Online First: Epub Date]|.
49. Aadahl M, Zacho M, Linneberg A, et al. Comparison of the Danish step test and the watt-max test for estimation of maximal oxygen uptake: the Health2008 study. European journal of preventive cardiology 2013;20(6):1088-94 doi: 10.1177/2047487312462825[published Online First: Epub Date]|.
50. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta psychiatrica Scandinavica 1983;67(6):361-70
51. Bjorngaard JH, Gunnell D, Elvestad MB, et al. The causal role of smoking in anxiety and depression: a Mendelian randomization analysis of the HUNT study. Psychological medicine 2013;43(4):711-9 doi: 10.1017/S0033291712001274[published Online First: Epub Date]|.
52. Jorgensen T, Borch-Johnsen K, Thomsen TF, et al. A randomized non-pharmacological intervention study for prevention of ischaemic heart disease: baseline results Inter99. European journal of cardiovascular prevention and rehabilitation : official journal of the European Society of Cardiology, Working Groups on Epidemiology & Prevention and Cardiac Rehabilitation and Exercise Physiology 2003;10(5):377-86 doi: 10.1097/01.hjr.0000096541.30533.82[published Online First: Epub Date]|.
53. Husemoen LL, Jorgensen T, Borch-Johnsen K, et al. The association of alcohol and alcohol metabolizing gene variants with diabetes and coronary heart disease risk factors in a white

population. PloS one 2010;5(8):e11735 doi: 10.1371/journal.pone.0011735[published Online First: Epub Date]|.
54. Jorgensen T, Ladelund S, Borch-Johnsen K, et al. Screening for risk of cardiovascular disease is not associated with mental distress: the Inter99 study. Preventive medicine 2009;48(3):242-6 doi: 10.1016/j.ypmed.2008.12.010[published Online First: Epub Date]|.
55. Rantakallio P. Groups at risk in low birth weight infants and perinatal mortality. Acta paediatrica Scandinavica 1969;193:Suppl 193:1+
56. Miettunen J, Isohanni M, Paunio T, et al. Predicting depression with psychopathology and temperament traits: the northern Finland 1966 birth cohort. Depression research and treatment 2012;2012:160905 doi: 10.1155/2012/160905[published Online First: Epub Date]|.
57. Sandanger I, Moum T, Ingebrigtsen G, et al. Concordance between symptom screening and diagnostic procedure: the Hopkins Symptom Checklist-25 and the Composite International Diagnostic Interview I. Social psychiatry and psychiatric epidemiology 1998;33(7):345-54
58. Jarvelin MR, Elliott P, Kleinschmidt I, et al. Ecological and individual predictors of birthweight in a northern Finland birth cohort 1986. Paediatric and perinatal epidemiology 1997;11(3):298-312
59. Achenbach TM. *Manual for the ASEBA School-Age Forms and Profiles*. University of Vermont, Research Center for Children, Youth and Families: Burlington, VT. , 1991.
60. Kantomaa MT, Tammelin TH, Demakakos P, et al. Physical activity, emotional and behavioural problems, maternal education and self-reported educational performance of adolescents. Health education research 2010;25(2):368-79 doi: 10.1093/her/cyp048[published Online First: Epub Date]|.
61. U.S. Department of Health and Human Services (DHHS). National Center for Health Statistics. NHANES III reference manuals and reports (CD-ROM). Hyattsville, MD: Centers for Disease Control and Prevention, 1996. Available from National Technical Information Service (NTIS), Springfield, VA. Acrobat .PDF format; includes access software: Adobe Systems, Inc. Acrobat Reader 2.1.
62. Woodruff R. A simple method for approximating the variance of a complicated estimate. . Journal of the American Staistical Association 1971;66:411-14
63. Chang MH, Lindegren ML, Butler MA, et al. Prevalence in the United States of selected candidate gene variants: Third National Health and Nutrition Examination Survey, 1991-1994. American journal of epidemiology 2009;169(1):54-66 doi: 10.1093/aje/kwn286[published Online First: Epub Date]|.
64. Wadsworth M, Kuh D, Richards M, et al. Cohort Profile: The 1946 National Birth Cohort (MRC National Survey of Health and Development). International journal of epidemiology 2006;35(1):49-54 doi: 10.1093/ije/dyi201[published Online First: Epub Date]|.
65. Kuh D, Pierce M, Adams J, et al. Cohort profile: updating the cohort profile for the MRC National Survey of Health and Development: a new clinic-based data collection for ageing research. International journal of epidemiology 2011;40(1):e1-9 doi: 10.1093/ije/dyq231[published Online First: Epub Date]|.
66. Rousseau K, Vinall LE, Butterworth SL, et al. MUC7 haplotype analysis: results from a longitudinal birth cohort support protective effect of the MUC7*5 allele on respiratory function. Ann Hum Genet 2006;70(Pt 4):417-27 doi: 10.1111/j.1469-1809.2006.00250.x[published Online First: Epub Date]|.
67. Hatch SL, Jones PB, Kuh D, et al. Childhood cognitive ability and adult mental health in the British 1946 birth cohort. Social science & medicine 2007;64(11):2285-96 doi: 10.1016/j.socscimed.2007.02.027[published Online First: Epub Date]|.
68. Clennell S, Kuh D, Guralnik JM, et al. Characterisation of smoking behaviour across the life course and its impact on decline in lung function and all-cause mortality: evidence from a British birth cohort. Journal of epidemiology and community health 2008;62(12):1051-6 doi: 10.1136/jech.2007.068312[published Online First: Epub Date]|.
69. Willemsen G, Vink JM, Abdellaoui A, et al. The Adult Netherlands Twin Register: twenty-five years of survey and biological data collection. Twin research and human genetics : the

official journal of the International Society for Twin Studies 2013;16(1):271-81 doi: 10.1017/thg.2012.140[published Online First: Epub Date]|.
70. Nivard MG, Mbarek H, Hottenga JJ, et al. Further confirmation of the association between anxiety and CTNND2: replication in humans. Genes, brain, and behavior 2014;13(2):195-201 doi: 10.1111/gbb.12095[published Online First: Epub Date]|.
71. Scheet P, Ehli EA, Xiao X, et al. Twins, tissue, and time: an assessment of SNPs and CNVs. Twin research and human genetics : the official journal of the International Society for Twin Studies 2012;15(6):737-45 doi: 10.1017/thg.2012.61[published Online First: Epub Date]|.
72. Willemsen G, de Geus EJ, Bartels M, et al. The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. Twin research and human genetics : the official journal of the International Society for Twin Studies 2010;13(3):231-45 doi: 10.1375/twin.13.3.231[published Online First: Epub Date]|.
73. Achenbach TM, Bernstein A, Dumenci L. DSM-oriented scales and statistically based syndromes for ages 18 to 59: linking taxonomic paradigms to facilitate multitaxonomic approaches. Journal of personality assessment 2005;84(1):49-63 doi: 10.1207/s15327752jpa8401_10[published Online First: Epub Date]|.
74. Johnstone EC, Yudkin PL, Hey K, et al. Genetic variation in dopaminergic pathways and short-term effectiveness of the nicotine patch. Pharmacogenetics 2004;14(2):83-90
75. Munafo MR, Clark TG, Roberts KH, et al. Neuroticism mediates the association of the serotonin transporter gene with lifetime major depression. Neuropsychobiology 2006;53(1):1-8 doi: 10.1159/000089915[published Online First: Epub Date]|.
76. Munafo MR, Johnstone EC, Walther D, et al. CHRNA3 rs1051730 genotype and short-term smoking cessation. Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco 2011;13(10):982-8 doi: 10.1093/ntr/ntr106[published Online First: Epub Date]|.
77. Hofman A, Breteler MM, van Duijn CM, et al. The Rotterdam Study: objectives and design update. European journal of epidemiology 2007;22(11):819-29 doi: 10.1007/s10654-007-9199-x[published Online First: Epub Date]|.
78. Wittchen HU, Lachner G, Wunderlich U, et al. Test-retest reliability of the computerized DSM-IV version of the Munich-Composite International Diagnostic Interview (M-CIDI). Social psychiatry and psychiatric epidemiology 1998;33(11):568-78
79. Association AP. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. Washington DC, 1994.
80. Pausova Z, Paus T, Abrahamowicz M, et al. Genes, maternal smoking, and the offspring brain and body during adolescence: design of the Saguenay Youth Study. Human brain mapping 2007;28(6):502-18 doi: 10.1002/hbm.20402[published Online First: Epub Date]|.
81. Melka MG, Gillis J, Bernard M, et al. FTO, obesity and the adolescent brain. Human molecular genetics 2013;22(5):1050-8 doi: 10.1093/hmg/dds504[published Online First: Epub Date]|.
82. Poulin C, Hand D, Boudreau B. Validity of a 12-item version of the CES-D used in the National Longitudinal Study of Children and Youth. Chronic diseases in Canada 2005;26(2-3):65-72
83. Marmot M, Brunner E. Cohort Profile: the Whitehall II study. International journal of epidemiology 2005;34(2):251-6 doi: 10.1093/ije/dyh372[published Online First: Epub Date]|.
84. Voight BF, Kang HM, Ding J, et al. The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. PLoS genetics 2012;8(8):e1002793 doi: 10.1371/journal.pgen.1002793[published Online First: Epub Date]|.
85. Stansfeld SA, Marmot MG. Social class and minor psychiatric disorder in British Civil Servants: a validated screening survey using the General Health Questionnaire. Psychological medicine 1992;22(3):739-49