medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Health-related effects of education level: a Mendelian randomization study

Shuai Yuan[1,2], Ying Xiong[3], Madeleine Michaëlsson[4], Karl Michaëlsson[1], Susanna C. Larsson[1,2]

Affiliations

[1] Department of Surgical Sciences, Uppsala University, Uppsala, Sweden

[2] Unit of Cardiovascular and Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

[3] Department of Public Health Sciences, Karolinska Institutet, Stockholm, Sweden

[4] Department of Education, Health and Social Studies, Dalarna University, Falun, Sweden

shuai.yuan@ki.se; ying.xiong@stud.ki.se; mmh@du.se

karl.michaelsson@surgsci.uu.se; susanna.larsson@surgsci.uu.se

Corresponding author:

Susanna C. Larsson, Department of Surgical Sciences, Uppsala University, Dag Hammarskjölds Väg 14B, Uppsala, 75185, Sweden; and Institute of Environmental Medicine, Karolinska Institutet, Nobelsväg 13, Stockholm, 17177, Sweden.

E-mail: susanna.larsson@surgsci.uu.se

Running head: Education and health outcomes

Word count in abstract / main test: 248 / 3795

Number of Tables / Figures: 2 / 2

Number of supplementary tables / figures: 3 / 6

Acknowledgments Summary-level data for education level, intelligence and outcomes were obtained from the UK Biobank and genetic consortia, including Psychiatric Genomics Consortium (bipolar disorder, schizophrenia, obsessive-compulsive disorder, post-traumatic stress disorder, anxiety and anorexia nervosa), The International Genomics of Alzheimer's Project (IGAP), Project MinE, Center for Neurogenomics and Cognitive Research (CNCR), iPSYCH, CARDIoGRAMplusC4D Consortium, MEGASTROKE Consortium, Atrial

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Fibrillation Consortium (AFGen), International Stroke Genetics Conosrtium (ISGC), Breast Cancer Association Consortium (BCAC), The Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) consortium, The International Lung Cancer Consortium (ILCCO), The Ovarian Cancer Association Consortium (OCAC), The DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) consortium, The Global Urate Genetics Consortium (GUGC), Genetics and Allied research in Rheumatic diseases Networking (GARNET) consortium, UK Inflammatory Bowel Disease Genetics (IBD) Consortium, CKDGen Consortium, the EArly Genetics and Lifecourse Epidemiology (EAGLE) Consortium, GEnetic Factors for OSteoporosis (GEFOS) Consortium, GIANT consortium, Neale lab, Consortium of Alcohol and Nicotine use (GSCAN). The authors thank all investigators for sharing these data. The list of investigators of MEGASTROKE is available at http://megastroke.org/authors.html. Funding of the MEGASTROKE project are specified at megastroke.org/acknowledgements.html.

Author contributions SY drafted the manuscript and conducted the statistical analyses. SY and YX conducted the systematic review. All authors contributed to the interpretation of the results and critical revision of the manuscript for important intellectual content and approved the final version of the manuscript.

Sources of funding Funding for this study came from the Swedish Research Council (Vetenskapsrådet; Grant Number 2019-00977), the Swedish Research Council for Health, Working Life and Welfare (Forte; Grant Number 2018-00123), and the Swedish Heart-Lung Foundation (Hjärt-Lungfonden; Grant number 20190247).

Conflicts of interest All authors declared no support from any organization for the submitted work; no financial relationships with any organization that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Ethical approval This MR study was approved by the Swedish Ethical Review Authority.

Informed consent All participants included in the genome-wide association studies gave informed consent.

Data sharing The datasets analyzed in this study are publicly available summary statistics.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Background

A deeper understanding of the causal links from education level to health outcomes may shed a light for disease prevention at a novel and efficient perspective.

Methods

We conducted a wide-angled Mendelian randomization to disentangle the causal role of education level from intelligence for 31 health outcomes and explore to what extent body mass index and smoking mediate the associations. Univariable and multivariable Mendelian randomization analyses were performed.

Results

Genetically higher education level was associated with lower risk of major mental disorders and most somatic diseases independent of intelligence. The intelligence-adjusted odds ratios for each additional standard deviation of education (4.2 years) were 0.48 (0.37, 0.62) for suicide attempts, 0.50 (0.36, 0.68) for large artery stroke, 0.51 (0.42, 0.63) for heart failure, 0.52 (0.42, 0.65) for lung cancer, 0.45 (0.33,0.61) for rheumatoid arthritis, and 0.48 (0.43, 0.55) for type 2 diabetes. Higher education level adjusted for intelligence was additionally associated with lower risk of insomnia, major depressive disorder, stroke, coronary artery disease, breast cancer, ovarian cancer and gout but with higher risk of obsessive-compulsive disorder, anorexia nervosa, bipolar disorder and prostate cancer. Moreover, higher education level was associated with modifiable health-related risk factors in a favorable manner. Adjustment for body mass index and smoking attenuated the associations between education level and several outcomes, especially for type 2 diabetes and heart failure. High education level exerts causal protective effects on major somatic diseases.

Conclusions

These findings emphasize the importance of education to reduce the burden of common diseases.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Keywords: education level; intelligence; Mendelian randomization analysis; somatic disorder; psychiatric disorder

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Main text

Education level is an important health social determinant and has been proposed as a modifiable risk factor for a number of disorders and diseases, such as depression [1], age-related cognitive decline [2], suicide [3], cardiovascular disease [4], cancer [5], and several other diseases [6-8]. However, it is unclear whether the associations are causal and independent of intelligence. Understanding the causal effects of education level on diseases can facilitate the aetiology pathway exploration of diseases as well as development of new strategies for disease prevention. Notwithstanding, randomized controlled trials are ethically and practically infeasible on this topic.

Exploiting genetic variants as instrumental variables for an exposure (e.g., education level), Mendelian randomization (MR) can strengthen the causal inference of an exposure-outcome association [9]. Comparing the risk of disease across individuals who have been classified by their genotype enables the causal effect of an exposure to be estimated with substantially less bias, such as confounding and reverse causality, than in a traditional observational analysis [9]. The rationale for diminished bias in MR studies is that genetic variants are randomly assorted and fixed at conception and therefore largely independent of confounders and cannot be modified by disease development [9].

We conducted an MR study to disentangle the causal role of education level from intelligence in major mental and neurological disorders and somatic diseases. A secondary aim was to explore whether intelligence is causally associated with the same health outcomes independently of education. We additionally investigated the associations of education level and intelligence with modifiable health-related risk factors and whether main lifestyle factors mediate the pathway from education to health outcomes.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Methods

Study design

The design and hypothesis of the present study are displayed in Supplementary Figure 1. We used summary-level data from large genome-wide association studies (GWASs) and genetic consortia (Table 1). Totally, our study included 11 mental and neurological disorders [10-20], 20 major somatic diseases [21-36], and 10 health-related risk factors [31,34,35,37-39]. A systematic review was conducted to find meta-analyses of observational studies of education level and diseases (Supplementary table 1).

Selection of instrumental variables

Instrumental variables for education level and intelligence were identified from GWASs of, respectively, 1 131 881 and 269 867 individuals of European ancestries [40,41]. In total, 1271 and 205 single-nucleotide polymorphisms (SNPs) at the genome-wide significance threshold ($p<5\times10^{-8}$) were proposed as instrumental variables for education level and intelligence. All used SNPs explained around 12.2% and 5.2% variance for education level and intelligence, respectively. Education level was defined as number of years of education and was unified across included studies according to an International Standard Classification of Education category. Across all cohorts, the sample-size-weighted mean of education year was 16.8 years of schooling with a standard deviation (SD) of 4.2 years. For the definition of intelligence as an instrumental variable in our analysis, included cohorts extracted a single sum score, mean score, or factor score from a multidimensional set of cognitive performance tests in GWAS with linear model, with the exception of High-IQ/Health and Retirement Study where a logistic regression GWAS was run with “case” status (high intelligence) versus controls (normal intelligence level). All included GWASs adjusted for key covariates,

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

such as age, sex and principal components for ancestry. The numbers of SNPs used as instrumental variables in each analysis are displayed in Table 1.

Outcome sources

Summary-level data for the associations of the education- and intelligence-associated SNPs with the outcomes were extracted from large-scale GWASs or genetic consortia. In the present MR study, we included 11 mental and neurological disorders [10-20], 9 cardiovascular diseases [21-25], 4 major cancers [26-29], 7 other diseases [30-36] and 10 established health-related risk factors for major diseases [31,34,35,37-39]. Detailed information, such as the number cases and controls, population structure and the source for each outcome, is presented in Table 1. Definitions of the diseases are presented in Supplementary Table 2.

Systematic review for meta-analysis of observational studies

A systematic literature search was conducted in the PubMed database before November 1st, 2019 to find meta-analyses of observational studies of education level in relation to diseases studied in the present MR study. We found latest published meta-analysis on 13 diseases and two risk factors, including major depressive disorders [42], suicide attempts [43], posttraumatic stress disorder [44], amyotrophic lateral sclerosis [45], Alzheimer's disease [46], coronary artery disease [47], heart failure [48], stroke [49], breast cancer [50], prostate cancer [51], lung cancer [52], type 2 diabetes [53], chronic kidney disease [54], body mass index [55] and hypertension (blood pressure) [56]. We extracted publication data (PMID number, the first author's name and year of publication), sample size, and risk estimates with their corresponding confidence intervals. Search strategy and characteristics of included meta-analyses are shown in Supplementary Table 1.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Statistical analysis

The random-effects inverse-variance weighted method was used to assess the associations of education and intelligence with the outcomes. The weighted median method and MR-Egger regression were used as sensitivity analyses to examine the consistency of results and to detect potential pleiotropy. The weighted median method gives accurate estimates if at least 50% of the instrumental variables are valid [57]. The MR-Egger regression can detect and adjust for pleiotropy albeit rendering low precision of the estimates [58]. Given the phenotypical and genetic correlation between education level and intelligence [59], we used a multivariable inverse-variance weighted method to disentangle the causal effect of education level on outcomes independent of intelligence and vice versa [60]. Considering that smoking and body mass index are modifiable risk factors linking education level to the most diseases, we also used multivariable-adjusted MR analysis to determine the effects of body mass index and smoking behavior on health outcomes for associations reaching the conventional significance level ($p<0.05$) in both univariable and intelligence-adjusted inverse-variance weighted model to explore the mediation effects of body mass index and smoking [61]. Proportions of attenuated effect size were calculated to present the magnitude of mediation effects. Odds ratios (ORs) and 95% confidence intervals (CIs) of diseases and changes of levels of risk factors were scaled to an SD increase in genetically predicted years of education (4.2 years) and intelligence. We calculated the power for the analyses of education level using a web-tool [62] and results are displayed in Table 1. All statistical analyses were two-sided and performed using the mrrobust package in Stata/SE 15.0 [63] and TwoSampleMR in R 3.6.0 software and MR-Base [64]. *P* values were not used strictly to define statistical significance; however, we interpreted the results based on the magnitude and strengths of the associations [65].

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Results

Genetically predicted education level and diseases

Genetically predicted education level was causally associated with most diseases, including 8 out of 11 mental and neurological disorders, all 9 studied cardiovascular diseases, all 4 studied cancers, and 4 out of 7 other common diseases in the univariable inverse-variance weighted MR analysis (Figure 1). In multivariable inverse-variance weighted analysis with adjustment for intelligence, higher education level was additionally associated with higher odds of schizophrenia and anxiety (these associations were not observed in the univariable analysis) (Figure 1). However, the inverse associations of education level with amyotrophic lateral sclerosis, Alzheimer's disease, cardioembolic stroke, intracerebral haemorrhage and inflammatory bowel disease observed in the crude MR analysis did not remain after adjustment for intelligence (Figure 1). In the intelligence-adjusted model, higher education level was associated with lower odds of rheumatoid arthritis (OR 0.45; 0.33, 0.61), type 2 diabetes (OR 0.48; 0.43, 0.55), suicide attempts (OR 0.48; 0.37, 0.62), large artery stroke (OR 0.50; 0.36, 0.68), heart failure (OR 0.51; 0.42, 0.63), lung cancer (OR 0.52; 0.42, 0.65), ovarian cancer (OR 0.53; 0.43, 0.66), small vessel stroke (OR 0.62; 0.47, 0.81), any ischemic stroke (OR 0.69; 0.61, 0.78), insomnia (OR 0.69; 0.64, 0.75), gout (OR 0.71; 0.60, 0.84), major depressive disorder (OR 0.78; 0.72, 0.85) and breast cancer (OR 0.85; 0.77, 0.94). Conversely, higher education level adjusted for intelligence was associated with an increased risk of obsessive-compulsive disorder (OR 2.24; 1.47, 3.41), bipolar disorder (OR 2.04; 1.64, 2.54), schizophrenia (OR 1.88; 1.49, 2.36), anorexia nervosa (OR 1.88; 1.53, 2.30) and anxiety (OR 1.84; 1.33, 2.56) (Figure 1). Results of sensitivity analyses were directionally similar but with wider CIs (Supplementary Figure 2). We detected possible pleiotropy in the analysis of obsessive-compulsive disorder, inflammatory bowel disease and rheumatoid arthritis (p for intercept in MR-Egger <0.05).

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Genetically predicted intelligence and diseases

The associations between intelligence and diseases are presented in Supplementary Figure 3 and 4. The effects of intelligence on mental and somatic diseases showed the same patterns with that of education level with the exception for schizophrenia (OR 0.47; 95% CI, 0.33, 0.67), bipolar disorder (OR 0.72; 95% CI, 0.52, 1.00), and Alzheimer's disease (OR 0.77, 95% CI, 0.59, 1.01) after adjustment for education level (Supplementary Figure 3). Similar results were obtained from the weighted median method and pleiotropy was detected in the analysis of heart failure in the MR-Egger regression analysis (Supplementary Figure 4).

Education, intelligence and risk factors

Genetically predicted higher education level was associated with later age of smoking initiation (β=0.31; 95% CI, 0.29, 0.33), higher bone mineral density (β=0.05; 95% CI, 0.03, 0.08), lower serum urate levels (β=-0.12; 95% CI, -0.15, -0.09), lower systolic (β=-0.13; 95% CI, -0.15, -0.11) and diastolic (β=-0.18; 95% CI, -0.20, -0.15) blood pressure, lower waist-to-hip ratio (β=-0.29; 95% CI, -0.31, -0.27), lower body mass index (β=-0.34; 95% CI, -0.37, -0.31) and fewer cigarettes smoked per day (β=-0.37; 95% CI, -0.42, -0.32) in the univariable model; the estimates were similar in the intelligence-adjusted model (Figure 2). Findings were consistent in sensitivity analyses and no pleiotropy was observed (Supplementary Figure 5). The same patterns of associations were observed for genetically predicted intelligence (Supplementary Figure 6). Nevertheless, after adjustment for education level, the magnitude of associations attenuated largely (Supplementary Figure 6).

Comparison with observational studies

The present MR findings were generally similar in the direction and magnitude to the estimates based on meta-analyses of observational studies (Supplementary Table 3). However,

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

there were discrepancies concerning the effects of education level on suicide attempts, breast cancer and prostate cancer.

Mediation effects of body mass index and smoking

Table 2 shows the results of mediation analyses after adjusting for body mass index and smoking behaviour in the pathway from education to health outcomes. Although not apparent for all disease outcomes, body mass index and smoking partly mediated most associations between education and diseases (Table 2). A strong mediation effect of body mass index was observed in the associations of education with gout (68%), type 2 diabetes (57%) and heart failure (34%). With regard to smoking, a strong mediation effect was detected in the association of education with major depressive disorder (24%), type 2 diabetes (20%) and lung cancer (19%). After adjustment for both body mass index and smoking, the direct causal effect of education on outcomes was substantially attenuated for gout (64%), type 2 diabetes (63%), heart failure (36%), obsessive-compulsive disorder (32%), suicide attempts (31%), coronary artery disease (31%) and lung cancer (29%).

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Discussion

In the present MR study with up to 1.3 million individuals, genetic predisposition to higher education level was causally associated with the majority of major health outcomes and related risk factors. Specifically, genetic predisposition to higher education level, independent of intelligence, was associated with lower risk of major depressive disorder, insomnia, suicide attempts, stroke, heart failure, coronary artery disease, breast cancer, ovarian cancer, lung cancer, gout, type 2 diabetes, and rheumatoid arthritis. Conversely, higher education level was associated with higher risk of obsessive-compulsive disorder, anorexia nervosa, bipolar disorder, and prostate cancer. Genetically predicted higher intelligence, independent of education, was inversely related to bipolar disorder, schizophrenia and Alzheimer's disease. Body mass index and smoking displayed strongest mediation effects observed for gout, type 2 diabetes, obsessive-compulsive disorder, suicide attempts and heart failure.

Comparison with previous studies

Our findings are broadly in line with a vast body of observational studies showing a protective association of high educational level on major depressive disorder [42], amyotrophic lateral sclerosis [45], Alzheimer's disease [46], coronary heart disease [47], heart failure [48], stroke [49], lung cancer [52], type 2 diabetes [53], chronic kidney disease [54], hypertension [56] and obesity [55]. However, for suicide attempts, posttraumatic stress disorder, breast cancer and prostate cancer, our MR findings differ from observational findings. The discrepancies might be attributed by reverse causality in the observational studies, heterogeneity and small sample sizes in the meta-analyses. A substantial heterogeneity (I^2=85%; p<0.001) was observed among included observational studies in the meta-analysis of breast cancer [50], and the sample size was small for prostate cancer [51]. Some studies have proposed that the higher risk of prostate cancer among men with high education level was driven by higher prostate-specific

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

antigen screening rate among educated men compared with men with low education level [66]. With regard to the inverse associations of higher education level with breast and ovarian cancer, these associations may in part be mediated by reproductive or hormone-related factors, or other health behaviours such as healthier diet and physical activity. We are not aware of any previous MR studies on education or intelligence in relation to prostate, breast, or ovarian cancer, but a protective causal effect of higher education on lung cancer risk has been reported recently [67].

Previous MR studies based on much fewer SNPs (up to 162) as predictors of education level showed a protective effect of higher educational level on coronary artery disease [68] and Alzheimer's disease [69]. The present study based on substantially more SNPs as instrumental variables more precisely verified the findings in previous research. Notably, the effects of high education level in the previous studies might be influenced by high intelligence given the tight phenotype and genetic correlation between intelligence and education level. In the present study, we used multivariable MR analysis to assess the direct effect of education level that is not mediated via intelligence with coronary artery disease and Alzheimer's disease. For Alzheimer's disease, we found that higher intelligence rather than education level may be the protective factor, whereas higher education level was the protective factor for coronary artery disease. In a previous MR study of the direct effect of education and intelligence on certain health outcomes, including diabetes, hypertension, heart attack, total stroke, total cancer, and depression, no significant association with education or intelligence was observed despite significant or suggestive associations of genetically predicted education with potential risk factors (blood pressure, smoking, alcohol consumption, body mass index, vigorous physical activity, and television watching) [59]. However, the genetic instrument used in that study comprised only 219 SNPs associated with education or intelligence and genetic associations with the outcomes were estimated in up to

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

138 670 UK Biobank participants, with relatively few outcomes. Thus, the power may have been insufficient to detect associations. Findings of other MR studies of education level in relation to health-related risk factors, such as obesity [70], cigarette smoking [71] and blood pressure [72], are consistent with our findings.

The inverse association between intelligence and schizophrenia was also observed in an observational study with 24 706 Swedish adults independent of overlapping genetic risks of two traits [73]. However, the effects of high intelligence on bipolar disorder were conflicting across observational studies. A large-scale cohort study with over 20-year follow-up duration proposed a "reversed-J" shape association between intelligence and risk of bipolar disorder, which indicated that individuals with the lowest and highest intelligence had the greatest risk of bipolar disorder) [74]. In a GWAS, it was revealed that bipolar disorder risk alleles were associated with better cognitive performance [75], which is opposite to our findings. Since education had a contradictory effect, which we observed in the present study, on bipolar disease to intelligence, the discrepancy among these studies might be caused by a mixed effect of education and intelligence at both aspects of phenotype and gene.

Possible mechanisms

Based on results of the present MR study and previous observational studies, there are three major possible pathways linking education level to health outcomes: 1) modifiable risk factors largely mediates the educational effects on diseases [72,76]; 2) there may be direct effects from education-related brain structures or function change via gene methylation, gene silencing etc. [77-79], especially for mental and neurological disorders; and 3) subjective well-being, happiness and meaning of life influenced by education level exerts effects on mental and somatic diseases directly or indirectly [80-83]. Education, as measured in this study, can be defined as an institutionalized form of social resource, and more specifically a form of

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

cultural capital drawing on the terminology of the French sociologist Pierre Bourdieu. Related forms of cultural capital emerge as objectivized resources – such as books, art or scientific tools – or incorporated resources, such as knowledge, attitudes and practices [84,85]. Our study shows that education is a health relevant cultural capital whilst intelligence is not to the same degree related with health and risk of disease.

Observational studies have found that the associations between education level and diseases attenuated largely after adjustment for health-related risk factors. Compared with unadjusted model, the risk of cardiovascular diseases of low education attainment attenuated around 30-45% in statistical models adjusted for multiple risk factors, such as smoking, body mass index, hypertension and physical activity [86,87]. However, measurement error and misclassification of mediators in observational studies often underestimates the mediation effects. The mediation effects were also proved in previous MR analysis [72,76]. In the present study, genetically predicted education level was associated with a favourable risk factor profile: with improved smoking behaviours (postponed smoking initiation age and less cigarette per day) as well as lower adiposity (body mass index and waist-to-hip ratio), blood pressure and serum urate levels, which might mediate associations between education level and diseases. By conducting mediation analysis, we showed that body mass index and smoking behaviour partly or entirely mediated the pathway from education level to several health outcomes, in particular gout, type 2 diabetes, obsessive-compulsive disorder, suicide attempts, atrial fibrillation, heart failure, coronary artery disease and lung cancer.

Previous studies have found that low education level might influence the changes in biochemical response and risk-related brain function, such as inflammation [77], cardiometabolic traits [78] and amygdala reactivity [79], via gene methylation, thereby influencing disease risk. In addition, genetic studies have also revealed that improvement of subjective well-being [80,81], happiness [80,81], meaning of life [82], social interaction [83], possibly derived from

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

high education level benefited human health directly and indirectly (e.g. influencing brain morphology, central nervous system and adrenal/pancreas tissues). There are other possible explanations, like followings: education level also could modify the risk of health outcomes through other diseases (comorbidity), the use of health care services, neighbourhood environment, occupations, income and marital status, which were amenable if education level was increased.

The results indicate that more than knowledge itself is affecting how people live their life, for instance through pathways regarding reduced smoking or alcohol habits among highly educated people. Therefore, we should consider further explanations, such as the relationship between high education on the one hand and the status and resources that follow it, on the other, which could by itself have a positive health effect on the individual. A further explanation assumes that it is the process itself that can be associated with increased well-being. That is, the process of taking part of and acquiring external knowledge rather than remaining with one's own innate thinking or being kept oblivious. Should only a fraction of the disease burden be explained by this process of mental activity given that education leads to a different kind of thinking, which is supported by the present study in that health is affected regardless of intelligence level then increased knowledge through education may lead to longevity through mechanisms beyond health literacy pathways of late-onset diseases and beyond the influence of social and material factors.

Strengths and limitations

The present study is the first study that comprehensively investigated the causal effects of education and intelligence on a very broad range of major disease outcomes and associated risk factors using genetic data from large-scale GWASs and genetic consortia. We used much more SNPs deriving from a larger GWAS with around 1.1 million individuals as instrumental

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

variables for education level and the latest GWASs with largest sample size for outcomes compared to previous MR studies, thereby assuring adequate statistical power to detect weak associations. In addition, we disentangled the independent effect of education level from intelligence using a multivariable MR approach. Thus, it is a straightforward approach to estimate the possible health benefits from education promotion among general population. We used mediation analysis to reveal the roles of body mass index and smoking behaviour as mediators in the pathway from education level to health outcomes. Even though there were genetic data for certain outcomes from GWASs with trans-ancestry populations, the majority of included participants were individuals with European ancestry thereby diminishing population stratification bias. However, population confinement limited the transferability of the present findings to populations of non-European ancestries.

The major limitation of the present study is the possible unbalanced horizontal pleiotropy aroused from used genetic variants marking more generic biological pathways. It has been found that the lead SNPs related to education level and intelligence are significantly overexpressed in the central nervous system, such as hippocampus and cerebral cortex, but not in other organs [40]. For cardiovascular disease, cancers and other physical diseases, we can minimize the possibility of pleiotropy from the global or systemic measures of fitness (such as mitochondrial function). It is more likely to conclude that the potential pleiotropy might exert a large to moderate effect via predominantly neurological pathways (for example, behaviours associated with obesity or smoking) for somatic diseases. In this scenario, the vertical pleiotropy would not bias the total causal effect by a higher educational level on disease development. With regard to mental and neurological disorders, although gene overwhelmingly expressed in the brain or central nervous system, studies found no, or at most a small, genetic correlation between lower education attainment and mental and neurological disorders by using bivariate genomic-relationship-matrix restricted maximum

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

likelihood analysis [1]. Thus, the associations between education level and mental or neurological diseases were not mainly because of measurable pleiotropic genetic effects, but because of education-related environmental factors. In addition, from a statistical perspective, we detected almost no pleiotropy in the results of MR-Egger regression and the estimates were consistent through sensitivity analyses, which indicated a negligible distortion by pleiotropy. Intergenerational effects from parents for certain disease, such as coronary artery diseases and type 2 diabetes, could not be assessed by using the data in the present MR study.

In summary, the present MR study strengthened the evidence of protective role of high education level on the majority of mental disorders and somatic diseases independent of intelligence. Body mass index and smoking partly mediated several of the associations between education level and health outcomes. These findings strongly suggest increasing education level for overall health benefits.

Table 1. Characteristics of included studies of mental disorders, somatic diseases, and health-related risk factors

Disease	Cases, No.	Controls, No.	Population	Used SNPs, No.	Source	Consortium
Mental disorder						
Anorexia nervosa	16 992	55 525	European	1227	Watson HJ et al	Psychiatric Genomics Consortium
Anxiety	7016	14 745	European	1207	Otowa T et al	Psychiatric Genomics Consortium
Bipolar disorder	20 352	31 358	European	1271	Stahl EA et al	Psychiatric Genomics Consortium
Insomnia	397 959	933 051	European	1194	Jansen PR et al	CNCR
Major depressive disorder	170 756	329 443	European	1244	Stahl EA et al	UK Biobank
Obsessive-compulsive disorder	2688	7037	European	1265	IOCDF-GC and OCGAS	Psychiatric Genomics Consortium
Posttraumatic stress disorder	30 000	170 000	Mix	1271	Nievergelt CM et al	Psychiatric Genomics Consortium
Suicide attempts	6024	44 240	European	1072	Erlangsen A et al	iPSYCH
Schizophrenia	33 426	54 065	European	1253	Psychiatric Genomics Consortium	Psychiatric Genomics Consortium
Neurological disease						
Amyotrophic lateral sclerosis	21 982	41 944	European	1268	Kunkle BW et al	Project MinE
Alzheimer's disease	12 577	23 475	European	1260	van Rheenen W et al	IGAP
Cardiovascular disease						
Atrial fibrillation	65 446	522 000	Mix	1261	Roselli C et al	AFGen
Coronary artery disease	60 801	123 504	Mix	1266	Nikpay M et al	CARDIoGRAMplusC4D Consortium
Heart failure	7382	387 652	European	1255	Aragam KG et al	UK Biobank
Total stroke	67 162	454 450	Mix	1258	Malik R et al	MEGASTROKE Consortium
Any ischemic stroke	60 341	NA	Mix	1270	Malik R et al	MEGASTROKE Consortium
Large artery stroke	6688	146 392	Mix	1270	Malik R et al	MEGASTROKE Consortium
Small vessel stroke	11 710	192 662	Mix	1268	Malik R et al	MEGASTROKE Consortium
Cardioembolic stroke	9006	204 570	Mix	1269	Malik R et al	MEGASTROKE Consortium
Intracerebral haemorrhage	1545	1481	Mix	852	Woo D et al	ISGC
Cancer						
Breast cancer	122 977	105 974	Mix	1269	Michailidou K et al	BCAC
Breast cancer ER+	69 501	NA	Mix	1269	Michailidou K et al	BCAC
Breast cancer ER-	21 468	NA	Mix	1269	Michailidou K et al	BCAC
Lung cancer	11 348	15 861	European	1230	Wang Y et al	ILCCO
Ovarian cancer	25 509	40 941	European	1213	Phelan CM et al	OCAC

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Prostate cancer	79 194	61 112	European	1175	Schumacher FR et al	PRACTICAL Consortium
Atopic dermatitis	21 399	95 464	Mix	1269	Paternoster L et al	EAGLE Consortium
Chronic kidney disease	41 395	439 303	European	1263	Wuttke M et al	CKDGen Consortium
Fracture	53 184	373 611	European	1259	Morris JA et al	GEFOS Consortium
Gout	13 179	750 634	Mix	1271	Tin A et al	GUGC
Inflammatory bowel disease	25 042	34 915	European	1213	de Lange KM et al	UK IBD consortium
Rheumatoid arthritis	29 880	73 758	Mix	1254	Okada Y et al	GARNET consortium
Type 2 diabetes	74 124	824 006	European	1263	Mahajan A et al	DIAGRAM consortium
Health-related risk factor						
Body mass index	NA	694 649	Mix	1260	Pulit SL et al	GIANT consortium
Waist-to-hip ratio	NA	697 734	Mix	1260	Pulit SL et al	GIANT consortium
Blood pressure, systolic	NA	317 754	European	1263	Neale lab	UK Biobank
Blood pressure, diastolic	NA	317 756	European	1263	Neale lab	UK Biobank
Age of smoking initiation	NA	341 427	European	1255	Liu M et al	GSCAN
Cigarettes per day	NA	337 334	European	1255	Liu M et al	GSCAN
Alcohol intake per week	NA	941 280	European	1255	Liu M et al	GSCAN
Bone mineral density	NA	426 824	European	1208	Morris JA et al	GEFOS Consortium
Estimated glomerular filtration rate	NA	765 348	Mix	1264	Wuttke M et al	CKDGen Consortium
Serum urate	NA	288 649	European	1264	Tin A et al	GUGC

AFGen stands for Atrial Fibrillation Consortium; BCAC, Breast Cancer Association Consortium; CNCR, Center for Neurogenomics and Cognitive Research; DIAGRAM, The DIAbetes Genetics Replication And Meta-analysis; EAGLE, The EArly Genetics and Lifecourse Epidemiology; ER, estrogen receptor; GARNET, Genetics and Allied research in Rheumatic diseases Networking; GEFOS, GEnetic Factors for Osteoporosis, GUGC, The Global Urate Genetics Consortium; GSCAN, Consortium of Alcohol and Nicotine use; IGAP, The International Genomics of Alzheimer's Project; ILCCO, The International Lung Cancer Consortium; ISGC, International Stroke Genetics Consortium; NA, Not available; OCAC, The Ovarian Cancer Association Consortium; PRACTICAL, The Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome; SNP, single-nucleotide polymorphism; UK IBD consortium, UK Inflammatory Bowel Disease Genetics Consortium.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Table 2. Mediation analysis to disentangle the effects of body mass index and smoking in the pathway from education level to health outcomes

Health outcome	Total effect of education			Effect after adjusting for BMI				Effect after adjusting for smoking				Effect after adjusting for both			
	OR[a]	95% CI	P	OR[b]	95% CI	P	%*	OR[c]	95% CI	P	%*	OR[d]	95% CI	P	%*
Mental disorder															
Anorexia nervosa	1.78	(1.60, 1.98)	1.56E-26	1.54	(1.36, 1.75)	1.70E-11	25	1.83	(1.63, 2.05)	3.00E-24	0	1.58	(1.39, 1.80)	4.90E-12	21
Insomnia	0.72	(0.69, 0.75)	4.15E-58	0.74	(0.71, 0.78)	2.80E-34	8	0.75	(0.72, 0.79)	1.10E-36	12	0.76	(0.73, 0.80)	1.80E-26	16
Major depressive disorder	0.77	(0.74, 0.80)	1.28E-36	0.80	(0.76, 0.84)	9.60E-20	15	0.82	(0.78, 0.85)	9.40E-19	24	0.83	(0.79, 0.87)	2.80E-13	2[illegible]
Obsessive-compulsive disorder	2.12	(1.71, 2.62)	5.66E-12	1.70	(1.31, 2.19)	5.50E-05	29	1.99	(1.57, 2.53)	1.10E-08	8	1.67	(1.28, 2.17)	2.00E-04	3[illegible]
Suicide attempts	0.56	(0.49, 0.64)	5.18E-18	0.64	(0.54, 0.75)	3.60E-08	23	0.61	(0.52, 0.70)	2.10E-11	15	0.67	(0.57, 0.79)	2.50E-06	3[illegible]
Neurological disease															
Amyotrophic lateral sclerosis	0.88	(0.79, 0.98)	1.50E-02	0.88	(0.78, 1.00)	4.60E-02	0	0.83	(0.74, 0.93)	1.00E-03	0	0.84	(0.74, 0.95)	8.00E-03	0
Alzheimer's disease	0.72	(0.66, 0.79)	6.60E-13	0.73	(0.65, 0.81)	4.17E-09	4	0.72	(0.65, 0.79)	3.00E-11	0	0.72	(0.64, 0.80)	5.01E-09	0
Cardiovascular disease															
Coronary artery disease	0.62	(0.58, 0.66)	8.54E-53	0.70	(0.65, 0.75)	2.30E-21	25	0.66	(0.61, 0.70)	9.30E-34	13	0.72	(0.67, 0.78)	2.50E-17	31
Heart failure	0.51	(0.46, 0.57)	4.16E-37	0.64	(0.56, 0.72)	3.50E-13	34	0.54	(0.48, 0.60)	1.80E-26	8	0.65	(0.57, 0.73)	1.20E-11	36
Total stroke	0.70	(0.67, 0.74)	1.10E-41	0.75	(0.71, 0.80)	1.10E-19	19	0.71	(0.67, 0.75)	5.40E-33	4	0.76	(0.71, 0.80)	3.70E-18	2[illegible]
Any ischemic stroke	0.68	(0.64, 0.72)	4.86E-33	0.70	(0.65, 0.76)	8.70E-20	8	0.68	(0.64, 0.73)	2.40E-27	0	0.70	(0.65, 0.76)	1.70E-18	8
Large artery stroke	0.51	(0.43, 0.60)	7.33E-17	0.55	(0.46, 0.67)	1.60E-09	11	0.53	(0.44, 0.63)	9.00E-13	6	0.56	(0.46, 0.69)	1.40E-08	14
Small vessel stroke	0.59	(0.51, 0.68)	3.07E-13	0.61	(0.52, 0.73)	1.40E-08	6	0.59	(0.50, 0.69)	2.00E-11	0	0.61	(0.51, 0.72)	2.60E-08	6
Cancer															
Breast cancer	0.87	(0.83, 0.91)	2.12E-08	0.88	(0.83, 0.94)	4.00E-05	8	0.88	(0.84, 0.93)	8.80E-06	8	0.89	(0.84, 0.95)	3.00E-04	1[illegible]
Breast cancer ER+	0.89	(0.84, 0.94)	2.34E-05	0.90	(0.84, 0.97)	4.00E-03	10	0.90	(0.85, 0.96)	1.00E-03	10	0.91	(0.85, 0.98)	1.20E-02	19
Breast cancer ER-	0.71	(0.66, 0.77)	6.97E-17	0.70	(0.64, 0.77)	1.80E-13	0	0.72	(0.66, 0.78)	4.20E-17	4	0.71	(0.64, 0.78)	2.50E-12	0
Lung cancer	0.52	(0.47, 0.58)	3.20E-30	0.58	(0.50, 0.66)	1.30E-15	17	0.59	(0.53, 0.67)	6.30E-17	19	0.63	(0.55, 0.72)	2.90E-11	29
Ovarian cancer	0.84	(0.78, 0.92)	7.30E-05	0.90	(0.81, 0.99)	3.60E-02	40	0.82	(0.75, 0.90)	4.00E-05	0	0.88	(0.79, 0.98)	1.40E-02	27
Prostate cancer	1.12	(1.05, 1.20)	1.00E-03	1.06	(0.98, 1.15)	1.45E-01	49	1.07	(1.00, 1.15)	6.00E-02	40	1.04	(0.96, 1.13)	3.23E-01	65

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Other diseases															
Gout	0.77	(0.70, 0.84)	1.75E-09	0.92	(0.84, 1.02)	1.35E-01	68	0.76	(0.69, 0.84)	4.40E-08	0	0.91	(0.81, 1.01)	6.70E-02	64
Rheumatoid arthritis	0.43	(0.37, 0.49)	5.29E-29	0.41	(0.34, 0.49)	1.50E-21	0	0.44	(0.37, 0.51)	7.90E-23	3	0.42	(0.35, 0.51)	1.40E-19	0
Type 2 diabetes	0.53	(0.50, 0.57)	9.07E-83	0.76	(0.71, 0.81)	1.20E-12	57	0.60	(0.56, 0.64)	3.30E-48	20	0.79	(0.74, 0.85)	6.30E-11	63

BMI indicates body mass index; ER, estrogen receptor.
[a] total effect without any adjustment;
[b] adjusted for the effect of body mass index;
[c] adjusted for the effect of smoking (cigarettes per day);
[d] adjusted for the effects of both body mass index and smoking behaviors;
*Percentage of the effect of education on the health outcome that is mediated by body mass index, smoking, or both (Formula: log(OR_total)-log(OR_adjusted)/log(OR_total)*100)). We replaced the values with zero for those percentage below zero.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 1

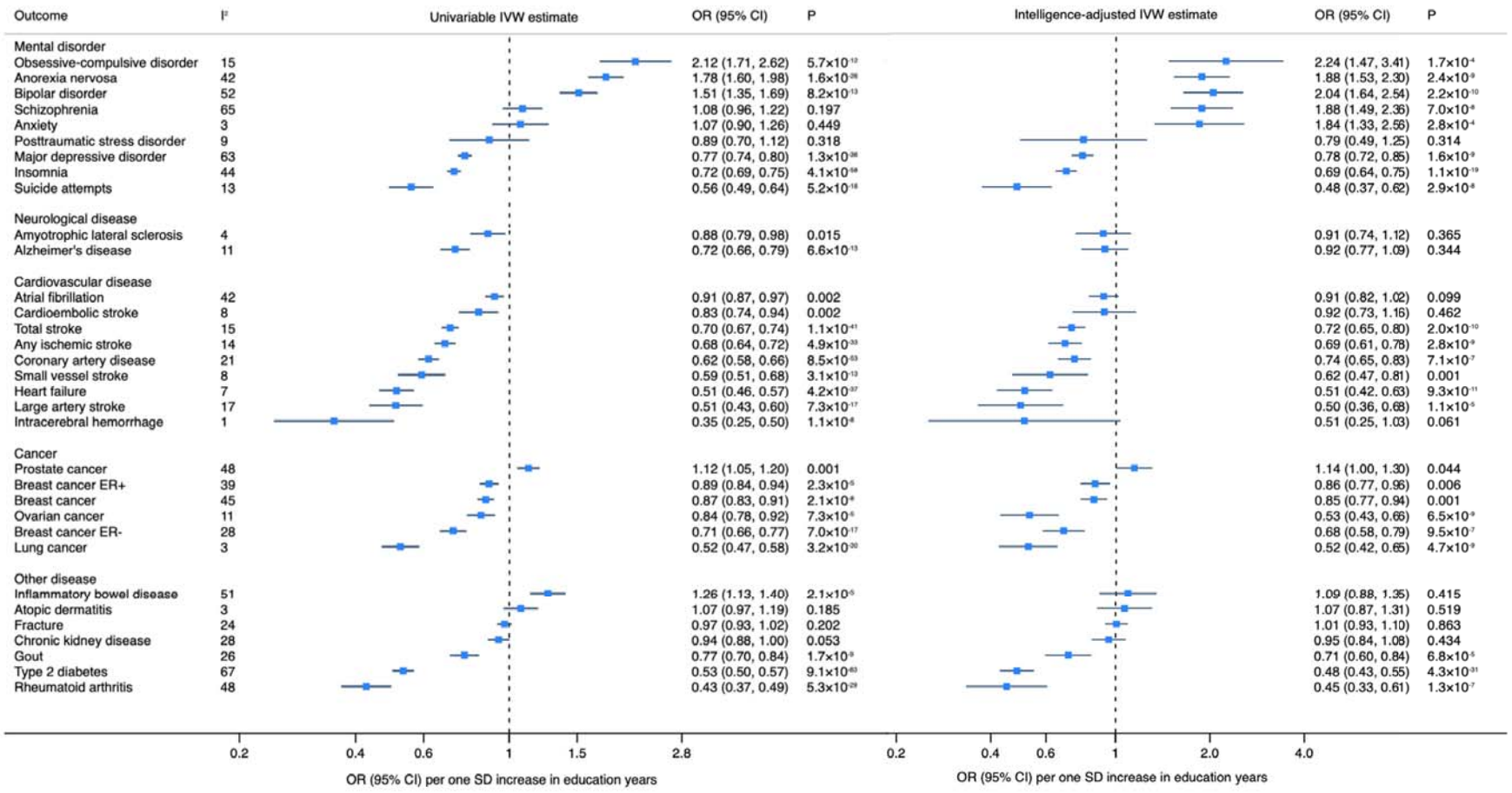
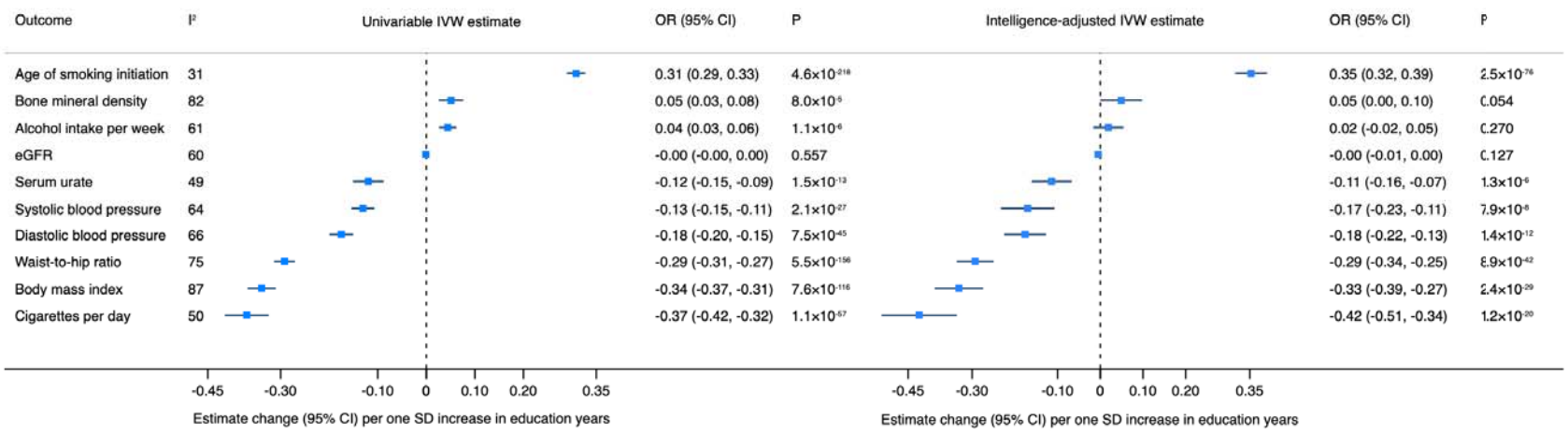


Figure 2



medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Figure legends

Figure 1. Associations of genetic predisposition to higher education level with health outcomes in MR analyses without and with adjustment for genetically predicted intelligence. CI indicates confidence interval; ER, estrogen receptor; IVW, inverse-variance weighted; OR, odds ratio; SD, standard deviation. I^2 represents the degree of heterogeneity among included SNPs for education.

Figure 2. Associations genetic predisposition to higher education level and health-related risk factors in MR analyses without and with adjustment for genetically predicted intelligence. CI indicates confidence interval; eGFR, estimated glomerular filtration rate; IVW, inverse-variance weighted; SD, standard deviation. I^2 represents the degree of heterogeneity among included SNPs for education.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

References

1. Peyrot WJ, Lee SH, Milaneschi Y, Abdellaoui A, Byrne EM, Esko T, de Geus EJ, Hemani G, Hottenga JJ, Kloiber S, Levinson DF, Lucae S, Martin NG, Medland SE, Metspalu A, Milani L, Noethen MM, Potash JB, Rietschel M, Rietveld CA, Ripke S, Shi J, Willemsen G, Zhu Z, Boomsma DI, Wray NR, Penninx BW. The association between lower educational attainment and depression owing to shared genetic effects? Results in ~25,000 subjects. *Mol Psychiatry* 2015;**20**(6):735-43.
2. Lipnicki DM, Crawford JD, Dutta R, Thalamuthu A, Kochan NA, Andrews G, Lima-Costa MF, Castro-Costa E, Brayne C, Matthews FE, Stephan BC, Lipton RB, Katz MJ, Ritchie K, Scali J, Ancelin ML, Scarmeas N, Yannakoulia M, Dardiotis E, Lam LC, Wong CH, Fung AW, Guaita A, Vaccaro R, Davin A, Kim KW, Han JW, Kim TH, Anstey KJ, Cherbuin N, Butterworth P, Scazufca M, Kumagai S, Chen S, Narazaki K, Ng TP, Gao Q, Reppermund S, Brodaty H, Lobo A, Lopez-Anton R, Santabarbara J, Sachdev PS. Age-related cognitive decline and associations with sex, education and apolipoprotein E genotype across ethnocultural groups and geographic regions: a collaborative cohort study. *PLoS Med* 2017;**14**(3):e1002261.
3. Naghavi M. Global, regional, and national burden of suicide mortality 1990 to 2016: systematic analysis for the Global Burden of Disease Study 2016. *Bmj* 2019;**364**:l94.
4. Rosengren A, Smyth A, Rangarajan S, Ramasundarahettige C, Bangdiwala SI, AlHabib KF, Avezum A, Bengtsson Bostrom K, Chifamba J, Gulec S, Gupta R, Igumbor EU, Iqbal R, Ismail N, Joseph P, Kaur M, Khatib R, Kruger IM, Lamelas P, Lanas F, Lear SA, Li W, Wang C, Quiang D, Wang Y, Lopez-Jaramillo P, Mohammadifard N, Mohan V, Mony PK, Poirier P, Srilatha S, Szuba A, Teo K, Wielgosz A, Yeates KE, Yusoff K, Yusuf R, Yusufali AH, Attaei MW, McKee M, Yusuf S. Socioeconomic status and risk of cardiovascular disease in 20 low-income, middle-income, and high-income countries: the Prospective Urban Rural Epidemiologic (PURE) study. *Lancet Glob Health* 2019;**7**(6):e748-e760.
5. Albano JD, Ward E, Jemal A, Anderson R, Cokkinides VE, Murray T, Henley J, Liff J, Thun MJ. Cancer mortality in the United States by education level and race. *J Natl Cancer Inst* 2007;**99**(18):1384-94.
6. Putrik P, Ramiro S, Keszei AP, Hmamouchi I, Dougados M, Uhlig T, Kvien TK, Boonen A. Lower education and living in countries with lower wealth are associated with higher disease activity in rheumatoid arthritis: results from the multinational COMORA study. *Ann Rheum Dis* 2016;**75**(3):540-6.
7. Patzer RE, McClellan WM. Influence of race, ethnicity and socioeconomic status on kidney disease. *Nat Rev Nephrol* 2012;**8**(9):533-41.
8. Mountjoy E, Davies NM, Plotnikov D, Smith GD, Rodriguez S, Williams CE, Guggenheim JA, Atan D. Education and myopia: assessing the direction of causality by mendelian randomisation. *Bmj* 2018;**361**:k2022.
9. Davey Smith G, Ebrahim S. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? *Bmj* 2005;**330**(7499):1076-9.
10. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, Mattheisen M, Wang Y, Coleman JRI, Gaspar HA, de Leeuw CA, Steinberg S, Pavlides JMW, Trzaskowski M, Byrne EM, Pers TH, Holmans PA, Richards AL, Abbott L, Agerbo E, Akil H, Albani D, Alliey-Rodriguez N, Als TD, Anjorin A, Antilla V, Awasthi S, Badner JA, Baekvad-Hansen M, Barchas JD, Bass N, Bauer M, Belliveau R, Bergen SE, Pedersen

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

CB, Boen E, Boks MP, Boocock J, Budde M, Bunney W, Burmeister M, Bybjerg-Grauholm J, Byerley W, Casas M, Cerrato F, Cervantes P, Chambert K, Charney AW, Chen D, Churchhouse C, Clarke TK, Coryell W, Craig DW, Cruceanu C, Curtis D, Czerski PM, Dale AM, de Jong S, Degenhardt F, Del-Favero J, DePaulo JR, Djurovic S, Dobbyn AL, Dumont A, Elvsashagen T, Escott-Price V, Fan CC, Fischer SB, Flickinger M, Foroud TM, Forty L, Frank J, Fraser C, Freimer NB, Frisen L, Gade K, Gage D, Garnham J, Giambartolomei C, Pedersen MG, Goldstein J, Gordon SD, Gordon-Smith K, Green EK, Green MJ, Greenwood TA, Grove J, Guan W, Guzman-Parra J, Hamshere ML, Hautzinger M, Heilbronner U, Herms S, Hipolito M, Hoffmann P, Holland D, Huckins L, Jamain S, Johnson JS, Jureus A, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet* 2019;**51**(5):793-803.

11. Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell* 2018;**173**(7):1705-1715.e16.
12. Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, Coleman JRI, Hagenaars SP, Ward J, Wigmore EM, Alloza C, Shen X, Barbu MC, Xu EY, Whalley HC, Marioni RE, Porteous DJ, Davies G, Deary IJ, Hemani G, Berger K, Teismann H, Rawal R, Arolt V, Baune BT, Dannlowski U, Domschke K, Tian C, Hinds DA, Trzaskowski M, Byrne EM, Ripke S, Smith DJ, Sullivan PF, Wray NR, Breen G, Lewis CM, McIntosh AM. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci* 2019;**22**(3):343-352.
13. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, Boland A, Vronskaya M, van der Lee SJ, Amlie-Wolf A, Bellenguez C, Frizatti A, Chouraki V, Martin ER, Sleegers K, Badarinarayan N, Jakobsdottir J, Hamilton-Nelson KL, Moreno-Grau S, Olaso R, Raybould R, Chen Y, Kuzma AB, Hiltunen M, Morgan T, Ahmad S, Vardarajan BN, Epelbaum J, Hoffmann P, Boada M, Beecham GW, Garnier JG, Harold D, Fitzpatrick AL, Valladares O, Moutet ML, Gerrish A, Smith AV, Qu L, Bacq D, Denning N, Jian X, Zhao Y, Del Zompo M, Fox NC, Choi SH, Mateo I, Hughes JT, Adams HH, Malamon J, Sanchez-Garcia F, Patel Y, Brody JA, Dombroski BA, Naranjo MCD, Daniilidou M, Eiriksdottir G, Mukherjee S, Wallon D, Uphill J, Aspelund T, Cantwell LB, Garzia F, Galimberti D, Hofer E, Butkiewicz M, Fin B, Scarpini E, Sarnowski C, Bush WS, Meslage S, Kornhuber J, White CC, Song Y, Barber RC, Engelborghs S, Sordon S, Voijnovic D, Adams PM, Vandenberghe R, Mayhaus M, Cupples LA, Albert MS, De Deyn PP, Gu W, Himali JJ, Beekly D, Squassina A, Hartmann AM, Orellana A, Blacker D, Rodriguez-Rodriguez E, Lovestone S, Garcia ME, Doody RS, Munoz-Fernadez C, Sussams R, Lin H, Fairchild TJ, Benito YA, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat Genet* 2019;**51**(3):414-430.
14. Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Mol Psychiatry* 2018;**23**(5):1181-1188.
15. van Rheenen W, Shatunov A, Dekker AM, McLaughlin RL, Diekstra FP, Pulit SL, van der Spek RA, Vosa U, de Jong S, Robinson MR, Yang J, Fogh I, van Doormaal PT, Tazelaar GH, Koppers M, Blokhuis AM, Sproviero W, Jones AR, Kenna KP, van Eijk KR, Harschnitz O, Schellevis RD, Brands WJ, Medic J, Menelaou A, Vajda A, Ticozzi N, Lin K, Rogelj B, Vrabec K, Ravnik-Glavac M, Koritnik B, Zidar J, Leonardis L, Groselj LD, Millecamps S, Salachas F, Meininger V, de Carvalho M, Pinto S, Mora JS, Rojas-Garcia R, Polak M, Chandran S, Colville S, Swingler R, Morrison KE, Shaw PJ, Hardy J, Orrell

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

RW, Pittman A, Sidle K, Fratta P, Malaspina A, Topp S, Petri S, Abdulla S, Drepper C, Sendtner M, Meyer T, Ophoff RA, Staats KA, Wiedau-Pazos M, Lomen-Hoerth C, Van Deerlin VM, Trojanowski JQ, Elman L, McCluskey L, Basak AN, Tunca C, Hamzeiy H, Parman Y, Meitinger T, Lichtner P, Radivojkov-Blagojevic M, Andres CR, Maurel C, Bensimon G, Landwehrmeyer B, Brice A, Payan CA, Saker-Delye S, Durr A, Wood NW, Tittmann L, Lieb W, Franke A, Rietschel M, Cichon S, Nothen MM, Amouyel P, Tzourio C, Dartigues JF, Uitterlinden AG, Rivadeneira F, Estrada K, Hofman A, Curtis C, Blauw HM, van der Kooi AJ, et al. Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nat Genet* 2016;**48**(9):1043-8.

16. Nievergelt CM, Maihofer AX, Klengel T, Atkinson EG, Chen CY, Choi KW, Coleman JRI, Dalvie S, Duncan LE, Gelernter J, Levey DF, Logue MW, Polimanti R, Provost AC, Ratanatharathorn A, Stein MB, Torres K, Aiello AE, Almli LM, Amstadter AB, Andersen SB, Andreassen OA, Arbisi PA, Ashley-Koch AE, Austin SB, Avdibegovic E, Babic D, Baekvad-Hansen M, Baker DG, Beckham JC, Bierut LJ, Bisson JI, Boks MP, Bolger EA, Borglum AD, Bradley B, Brashear M, Breen G, Bryant RA, Bustamante AC, Bybjerg-Grauholm J, Calabrese JR, Caldas-de-Almeida JM, Dale AM, Daly MJ, Daskalakis NP, Deckert J, Delahanty DL, Dennis MF, Disner SG, Domschke K, Dzubur-Kulenovic A, Erbes CR, Evans A, Farrer LA, Feeny NC, Flory JD, Forbes D, Franz CE, Galea S, Garrett ME, Gelaye B, Geuze E, Gillespie C, Uka AG, Gordon SD, Guffanti G, Hammamieh R, Harnal S, Hauser MA, Heath AC, Hemmings SMJ, Hougaard DM, Jakovljevic M, Jett M, Johnson EO, Jones I, Jovanovic T, Qin XJ, Junglen AG, Karstoft KI, Kaufman ML, Kessler RC, Khan A, Kimbrel NA, King AP, Koen N, Kranzler HR, Kremen WS, Lawford BR, Lebois LAM, Lewis CE, Linnstaedt SD, Lori A, Lugonja B, Luykx JJ, Lyons MJ, Maples-Keller J, Marmar C, Martin AR, et al. International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. *Nat Commun* 2019;**10**(1):4558.

17. Otowa T, Hek K, Lee M, Byrne EM, Mirza SS, Nivard MG, Bigdeli T, Aggen SH, Adkins D, Wolen A, Fanous A, Keller MC, Castelao E, Kutalik Z, Van der Auwera S, Homuth G, Nauck M, Teumer A, Milaneschi Y, Hottenga JJ, Direk N, Hofman A, Uitterlinden A, Mulder CL, Henders AK, Medland SE, Gordon S, Heath AC, Madden PA, Pergadia ML, van der Most PJ, Nolte IM, van Oort FV, Hartman CA, Oldehinkel AJ, Preisig M, Grabe HJ, Middeldorp CM, Penninx BW, Boomsma D, Martin NG, Montgomery G, Maher BS, van den Oord EJ, Wray NR, Tiemeier H, Hettema JM. Meta-analysis of genome-wide association studies of anxiety disorders. *Mol Psychiatry* 2016;**21**(10):1391-9.

18. Jansen PR, Watanabe K, Stringer S, Skene N, Bryois J, Hammerschlag AR, de Leeuw CA, Benjamins JS, Munoz-Manchado AB, Nagel M, Savage JE, Tiemeier H, White T, Tung JY, Hinds DA, Vacic V, Wang X, Sullivan PF, van der Sluis S, Polderman TJC, Smit AB, Hjerling-Leffler J, Van Someren EJW, Posthuma D. Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat Genet* 2019;**51**(3):394-403.

19. Watson HJ, Yilmaz Z, Thornton LM, Hubel C, Coleman JRI, Gaspar HA, Bryois J, Hinney A, Leppa VM, Mattheisen M, Medland SE, Ripke S, Yao S, Giusti-Rodriguez P, Hanscombe KB, Purves KL, Adan RAH, Alfredsson L, Ando T, Andreassen OA, Baker JH, Berrettini WH, Boehm I, Boni C, Perica VB, Buehren K, Burghardt R, Cassina M, Cichon S, Clementi M, Cone RD, Courtet P, Crow S, Crowley JJ, Danner UN, Davis OSP, de Zwaan M, Dedoussis G, Degortes D, DeSocio JE, Dick DM, Dikeos D, Dina C,

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Dmitrzak-Weglarz M, Docampo E, Duncan LE, Egberts K, Ehrlich S, Escaramis G, Esko T, Estivill X, Farmer A, Favaro A, Fernandez-Aranda F, Fichter MM, Fischer K, Focker M, Foretova L, Forstner AJ, Forzan M, Franklin CS, Gallinger S, Giegling I, Giuranna J, Gonidakis F, Gorwood P, Mayora MG, Guillaume S, Guo Y, Hakonarson H, Hatzikotoulas K, Hauser J, Hebebrand J, Helder SG, Herms S, Herpertz-Dahlmann B, Herzog W, Huckins LM, Hudson JI, Imgart H, Inoko H, Janout V, Jimenez-Murcia S, Julia A, Kalsi G, Kaminska D, Kaprio J, Karhunen L, Karwautz A, Kas MJH, Kennedy JL, Keski-Rahkonen A, Kiezebrink K, Kim YR, Klareskog L, Klump KL, Knudsen GPS, La Via MC, Le Hellard S, Levitan RD, et al. Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nat Genet* 2019;**51**(8):1207-1214.

20. Erlangsen A, Appadurai V, Wang Y, Turecki G, Mors O, Werge T, Mortensen PB, Starnawska A, Borglum AD, Schork A, Nudel R, Baekvad-Hansen M, Bybjerg-Grauholm J, Hougaard DM, Thompson WK, Nordentoft M, Agerbo E. Genetics of suicide attempts in individuals with and without mental disorders: a population-based genome-wide association study. *Mol Psychiatry* 2018.

21. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro K, Bjonnes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang SJ, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lyytikainen LP, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith AV, Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, de Andrade M, de Vries PS, van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han BG, Jalilzadeh S, Kessler T, Konig IR, Lannfelt L, Lieb W, Lind L, Lindgren CM, Lokki ML, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon FU, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, Stirrups KE, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples LA, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T, Feitosa MF, Franco OH, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015;**47**(10):1121-1130.

22. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, Rutten-Jacobs L, Giese AK, van der Laan SW, Gretarsdottir S, Anderson CD, Chong M, Adams HHH, Ago T, Almgren P, Amouyel P, Ay H, Bartz TM, Benavente OR, Bevan S, Boncoraglio GB, Brown RD, Jr., Butterworth AS, Carrera C, Carty CL, Chasman DI, Chen WM, Cole JW, Correa A, Cotlarciuc I, Cruchaga C, Danesh J, de Bakker PIW, DeStefano AL, den Hoed M, Duan Q, Engelter ST, Falcone GJ, Gottesman RF, Grewal RP, Gudnason V, Gustafsson S, Haessler J, Harris TB, Hassan A, Havulinna AS, Heckbert SR, Holliday EG, Howard G, Hsu FC, Hyacinth HI, Ikram MA, Ingelsson E, Irvin MR, Jian X, Jimenez-Conde J, Johnson JA, Jukema JW, Kanai M, Keene KL, Kissela BM, Kleindorfer DO, Kooperberg C, Kubo M, Lange LA, Langefeld CD, Langenberg C, Launer LJ, Lee JM, Lemmens R, Leys D, Lewis CM, Lin WY, Lindgren AG, Lorentzen E, Magnusson PK, Maguire J, Manichaikul A, McArdle PF, Meschia JF, Mitchell BD, Mosley TH, Nalls MA, Ninomiya T, O'Donnell MJ, Psaty BM, Pulit SL, Rannikmae K, Reiner AP, Rexrode KM, Rice K, Rich SS, Ridker PM, Rost NS, Rothwell PM, Rotter JI, Rundek T, Sacco RL, Sakaue S, Sale MM, et al. Multiancestry genome-wide association study of 520,000

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet* 2018;**50**(4):524-537.

23. Aragam KG, Chaffin M, Levinson RT, McDermott G, Choi SH, Shoemaker MB, Haas ME, Weng LC, Lindsay ME, Smith JG, Newton-Cheh C, Roden DM, London B, Wells QS, Ellinor PT, Kathiresan S, Lubitz SA. Phenotypic Refinement of Heart Failure in a National Biobank Facilitates Genetic Discovery. *Circulation* 2018.
24. Roselli C, Chaffin MD, Weng LC, Aeschbacher S, Ahlberg G, Albert CM, Almgren P, Alonso A, Anderson CD, Aragam KG, Arking DE, Barnard J, Bartz TM, Benjamin EJ, Bihlmeyer NA, Bis JC, Bloom HL, Boerwinkle E, Bottinger EB, Brody JA, Calkins H, Campbell A, Cappola TP, Carlquist J, Chasman DI, Chen LY, Chen YI, Choi EK, Choi SH, Christophersen IE, Chung MK, Cole JW, Conen D, Cook J, Crijns HJ, Cutler MJ, Damrauer SM, Daniels BR, Darbar D, Delgado G, Denny JC, Dichgans M, Dorr M, Dudink EA, Dudley SC, Esa N, Esko T, Eskola M, Fatkin D, Felix SB, Ford I, Franco OH, Geelhoed B, Grewal RP, Gudnason V, Guo X, Gupta N, Gustafsson S, Gutmann R, Hamsten A, Harris TB, Hayward C, Heckbert SR, Hernesniemi J, Hocking LJ, Hofman A, Horimoto A, Huang J, Huang PL, Huffman J, Ingelsson E, Ipek EG, Ito K, Jimenez-Conde J, Johnson R, Jukema JW, Kaab S, Kahonen M, Kamatani Y, Kane JP, Kastrati A, Kathiresan S, Katschnig-Winter P, Kavousi M, Kessler T, Kietselaer BL, Kirchhof P, Kleber ME, Knight S, Krieger JE, Kubo M, Launer LJ, Laurikka J, Lehtimaki T, Leineweber K, Lemaitre RN, Li M, Lim HE, Lin HJ, Lin H, et al. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet* 2018;**50**(9):1225-1233.
25. Woo D, Falcone GJ, Devan WJ, Brown WM, Biffi A, Howard TD, Anderson CD, Brouwers HB, Valant V, Battey TW, Radmanesh F, Raffeld MR, Baedorf-Kassis S, Deka R, Woo JG, Martin LJ, Haverbusch M, Moomaw CJ, Sun G, Broderick JP, Flaherty ML, Martini SR, Kleindorfer DO, Kissela B, Comeau ME, Jagiella JM, Schmidt H, Freudenberger P, Pichler A, Enzinger C, Hansen BM, Norrving B, Jimenez-Conde J, Giralt-Steinhauer E, Elosua R, Cuadrado-Godia E, Soriano C, Roquer J, Kraft P, Ayres AM, Schwab K, McCauley JL, Pera J, Urbanik A, Rost NS, Goldstein JN, Viswanathan A, Stogerer EM, Tirschwell DL, Selim M, Brown DL, Silliman SL, Worrall BB, Meschia JF, Kidwell CS, Montaner J, Fernandez-Cadenas I, Delgado P, Malik R, Dichgans M, Greenberg SM, Rothwell PM, Lindgren A, Slowik A, Schmidt R, Langefeld CD, Rosand J. Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. *Am J Hum Genet* 2014;**94**(4):511-21.
26. Michailidou K, Lindstrom S, Dennis J, Beesley J, Hui S, Kar S, Lemacon A, Soucy P, Glubb D, Rostamianfar A, Bolla MK, Wang Q, Tyrer J, Dicks E, Lee A, Wang Z, Allen J, Keeman R, Eilber U, French JD, Qing Chen X, Fachal L, McCue K, McCart Reed AE, Ghoussaini M, Carroll JS, Jiang X, Finucane H, Adams M, Adank MA, Ahsan H, Aittomaki K, Anton-Culver H, Antonenkova NN, Arndt V, Aronson KJ, Arun B, Auer PL, Bacot F, Barrdahl M, Baynes C, Beckmann MW, Behrens S, Benitez J, Bermisheva M, Bernstein L, Blomqvist C, Bogdanova NV, Bojesen SE, Bonanni B, Borresen-Dale AL, Brand JS, Brauch H, Brennan P, Brenner H, Brinton L, Broberg P, Brock IW, Broeks A, Brooks-Wilson A, Brucker SY, Bruning T, Burwinkel B, Butterbach K, Cai Q, Cai H, Caldes T, Canzian F, Carracedo A, Carter BD, Castelao JE, Chan TL, David Cheng TY, Seng Chia K, Choi JY, Christiansen H, Clarke CL, Collee M, Conroy DM, Cordina-Duverger E, Cornelissen S, Cox DG, Cox A, Cross SS, Cunningham JM, Czene K, Daly MB, Devilee P, Doheny KF, Dork T, Dos-Santos-Silva I, Dumont M, Durcan L, Dwek M,

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Eccles DM, Ekici AB, Eliassen AH, Ellberg C, Elvira M, Engel C, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature* 2017;**551**(7678):92-94.

27. Schumacher FR, Al Olama AA, Berndt SI, Benlloch S, Ahmed M, Saunders EJ, Dadaev T, Leongamornlert D, Anokian E, Cieza-Borrella C, Goh C, Brook MN, Sheng X, Fachal L, Dennis J, Tyrer J, Muir K, Lophatananon A, Stevens VL, Gapstur SM, Carter BD, Tangen CM, Goodman PJ, Thompson IM, Jr., Batra J, Chambers S, Moya L, Clements J, Horvath L, Tilley W, Risbridger GP, Gronberg H, Aly M, Nordstrom T, Pharoah P, Pashayan N, Schleutker J, Tammela TLJ, Sipeky C, Auvinen A, Albanes D, Weinstein S, Wolk A, Hakansson N, West CML, Dunning AM, Burnet N, Mucci LA, Giovannucci E, Andriole GL, Cussenot O, Cancel-Tassin G, Koutros S, Beane Freeman LE, Sorensen KD, Orntoft TF, Borre M, Maehle L, Grindedal EM, Neal DE, Donovan JL, Hamdy FC, Martin RM, Travis RC, Key TJ, Hamilton RJ, Fleshner NE, Finelli A, Ingles SA, Stern MC, Rosenstein BS, Kerns SL, Ostrer H, Lu YJ, Zhang HW, Feng N, Mao X, Guo X, Wang G, Sun Z, Giles GG, Southey MC, MacInnis RJ, FitzGerald LM, Kibel AS, Drake BF, Vega A, Gomez-Caamano A, Szulkin R, Eklund M, Kogevinas M, Llorca J, Castano-Vinyals G, Penney KL, Stampfer M, Park JY, Sellers TA, Lin HY, Stanford JL, Cybulski C, et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat Genet* 2018;**50**(7):928-936.

28. Wang Y, McKay JD, Rafnar T, Wang Z, Timofeeva MN, Broderick P, Zong X, Laplana M, Wei Y, Han Y, Lloyd A, Delahaye-Sourdeix M, Chubb D, Gaborieau V, Wheeler W, Chatterjee N, Thorleifsson G, Sulem P, Liu G, Kaaks R, Henrion M, Kinnersley B, Vallee M, LeCalvez-Kelm F, Stevens VL, Gapstur SM, Chen WV, Zaridze D, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates D, Bencko V, Foretova L, Janout V, Krokan HE, Gabrielsen ME, Skorpen F, Vatten L, Njolstad I, Chen C, Goodman G, Benhamou S, Vooder T, Valk K, Nelis M, Metspalu A, Lener M, Lubinski J, Johansson M, Vineis P, Agudo A, Clavel-Chapelon F, Bueno-de-Mesquita HB, Trichopoulos D, Khaw KT, Weiderpass E, Tjonneland A, Riboli E, Lathrop M, Scelo G, Albanes D, Caporaso NE, Ye Y, Gu J, Wu X, Spitz MR, Dienemann H, Rosenberger A, Su L, Matakidou A, Eisen T, Stefansson K, Risch A, Chanock SJ, Christiani DC, Hung RJ, Brennan P, Landi MT, Houlston RS, Amos CI. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. *Nat Genet* 2014;**46**(7):736-41.

29. Phelan CM, Kuchenbaecker KB, Tyrer JP, Kar SP, Lawrenson K, Winham SJ, Dennis J, Pirie A, Riggan MJ, Chornokur G, Earp MA, Lyra PC, Jr., Lee JM, Coetzee S, Beesley J, McGuffog L, Soucy P, Dicks E, Lee A, Barrowdale D, Lecarpentier J, Leslie G, Aalfs CM, Aben KKH, Adams M, Adlard J, Andrulis IL, Anton-Culver H, Antonenkova N, Aravantinos G, Arnold N, Arun BK, Arver B, Azzollini J, Balmana J, Banerjee SN, Barjhoux L, Barkardottir RB, Bean Y, Beckmann MW, Beeghly-Fadiel A, Benitez J, Bermisheva M, Bernardini MQ, Birrer MJ, Bjorge L, Black A, Blankstein K, Blok MJ, Bodelon C, Bogdanova N, Bojesen A, Bonanni B, Borg A, Bradbury AR, Brenton JD, Brewer C, Brinton L, Broberg P, Brooks-Wilson A, Bruinsma F, Brunet J, Buecher B, Butzow R, Buys SS, Caldes T, Caligo MA, Campbell I, Cannioto R, Carney ME, Cescon T, Chan SB, Chang-Claude J, Chanock S, Chen XQ, Chiew YE, Chiquette J, Chung WK, Claes KBM, Conner T, Cook LS, Cook J, Cramer DW, Cunningham JM, D'Aloisio AA, Daly MB, Damiola F, Damirovna SD, Dansonka-Mieszkowska A, Dao F, Davidson R, DeFazio A, Delnatte C, Doheny KF, Diez O, Ding YC, Doherty JA, Domchek SM, Dorfling CM, Dork T, et al. Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer. *Nat Genet* 2017;**49**(5):680-691.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

30. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, Payne AJ, Steinthorsdottir V, Scott RA, Grarup N, Cook JP, Schmidt EM, Wuttke M, Sarnowski C, Magi R, Nano J, Gieger C, Trompet S, Lecoeur C, Preuss MH, Prins BP, Guo X, Bielak LF, Below JE, Bowden DW, Chambers JC, Kim YJ, Ng MCY, Petty LE, Sim X, Zhang W, Bennett AJ, Bork-Jensen J, Brummett CM, Canouil M, Ec Kardt KU, Fischer K, Kardia SLR, Kronenberg F, Lall K, Liu CT, Locke AE, Luan J, Ntalla I, Nylander V, Schonherr S, Schurmann C, Yengo L, Bottinger EP, Brandslund I, Christensen C, Dedoussis G, Florez JC, Ford I, Franco OH, Frayling TM, Giedraitis V, Hackinger S, Hattersley AT, Herder C, Ikram MA, Ingelsson M, Jorgensen ME, Jorgensen T, Kriebel J, Kuusisto J, Ligthart S, Lindgren CM, Linneberg A, Lyssenko V, Mamakou V, Meitinger T, Mohlke KL, Morris AD, Nadkarni G, Pankow JS, Peters A, Sattar N, Stancakova A, Strauch K, Taylor KD, Thorand B, Thorleifsson G, Thorsteinsdottir U, Tuomilehto J, Witte DR, Dupuis J, Peyser PA, Zeggini E, Loos RJF, Froguel P, Ingelsson E, Lind L, Groop L, Laakso M, Collins FS, Jukema JW, Palmer CNA, Grallert H, Metspalu A, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet* 2018;**50**(11):1505-1513.

31. Tin A, Marten J, Halperin Kuhns VL, Li Y, Wuttke M, Kirsten H, Sieber KB, Qiu C, Gorski M, Yu Z, Giri A, Sveinbjornsson G, Li M, Chu AY, Hoppmann A, O'Connor LJ, Prins B, Nutile T, Noce D, Akiyama M, Cocca M, Ghasemi S, van der Most PJ, Horn K, Xu Y, Fuchsberger C, Sedaghat S, Afaq S, Amin N, Arnlov J, Bakker SJL, Bansal N, Baptista D, Bergmann S, Biggs ML, Biino G, Boerwinkle E, Bottinger EP, Boutin TS, Brumat M, Burkhardt R, Campana E, Campbell A, Campbell H, Carroll RJ, Catamo E, Chambers JC, Ciullo M, Concas MP, Coresh J, Corre T, Cusi D, Felicita SC, de Borst MH, De Grandi A, de Mutsert R, de Vries APJ, Delgado G, Demirkan A, Devuyst O, Dittrich K, Eckardt KU, Ehret G, Endlich K, Evans MK, Gansevoort RT, Gasparini P, Giedraitis V, Gieger C, Girotto G, Gogele M, Gordon SD, Gudbjartsson DF, Gudnason V, Haller T, Hamet P, Harris TB, Hayward C, Hicks AA, Hofer E, Holm H, Huang W, Hutri-Kahonen N, Hwang SJ, Ikram MA, Lewis RM, Ingelsson E, Jakobsdottir J, Jonsdottir I, Jonsson H, Joshi PK, Josyula NS, Jung B, Kahonen M, Kamatani Y, Kanai M, Kerr SM, Kiess W, Kleber ME, Koenig W, et al. Target genes, variants, tissues and transcriptional pathways influencing human serum urate levels. *Nat Genet* 2019.

32. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, Kochi Y, Ohmura K, Suzuki A, Yoshida S, Graham RR, Manoharan A, Ortmann W, Bhangale T, Denny JC, Carroll RJ, Eyler AE, Greenberg JD, Kremer JM, Pappas DA, Jiang L, Yin J, Ye L, Su DF, Yang J, Xie G, Keystone E, Westra HJ, Esko T, Metspalu A, Zhou X, Gupta N, Mirel D, Stahl EA, Diogo D, Cui J, Liao K, Guo MH, Myouzen K, Kawaguchi T, Coenen MJ, van Riel PL, van de Laar MA, Guchelaar HJ, Huizinga TW, Dieude P, Mariette X, Bridges SL, Jr., Zhernakova A, Toes RE, Tak PP, Miceli-Richard C, Bang SY, Lee HS, Martin J, Gonzalez-Gay MA, Rodriguez-Rodriguez L, Rantapaa-Dahlqvist S, Arlestig L, Choi HK, Kamatani Y, Galan P, Lathrop M, Eyre S, Bowes J, Barton A, de Vries N, Moreland LW, Criswell LA, Karlson EW, Taniguchi A, Yamada R, Kubo M, Liu JS, Bae SC, Worthington J, Padyukov L, Klareskog L, Gregersen PK, Raychaudhuri S, Stranger BE, De Jager PL, Franke L, Visscher PM, Brown MA, Yamanaka H, Mimori T, Takahashi A, Xu H, Behrens TW, Siminovitch KA, Momohara S, Matsuda F, Yamamoto K, Plenge RM. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 2014;**506**(7488):376-81.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

33. de Lange KM, Moutsianas L, Lee JC, Lamb CA, Luo Y, Kennedy NA, Jostins L, Rice DL, Gutierrez-Achury J, Ji SG, Heap G, Nimmo ER, Edwards C, Henderson P, Mowat C, Sanderson J, Satsangi J, Simmons A, Wilson DC, Tremelling M, Hart A, Mathew CG, Newman WG, Parkes M, Lees CW, Uhlig H, Hawkey C, Prescott NJ, Ahmad T, Mansfield JC, Anderson CA, Barrett JC. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* 2017;**49**(2):256-261.

34. Morris JA, Kemp JP, Youlten SE, Laurent L, Logan JG, Chai RC, Vulpescu NA, Forgetta V, Kleinman A, Mohanty ST, Sergio CM, Quinn J, Nguyen-Yamamoto L, Luco AL, Vijay J, Simon MM, Pramatarova A, Medina-Gomez C, Trajanoska K, Ghirardello EJ, Butterfield NC, Curry KF, Leitch VD, Sparkes PC, Adoum AT, Mannan NS, Komla-Ebri DSK, Pollard AS, Dewhurst HF, Hassall TAD, Beltejar MG, Adams DJ, Vaillancourt SM, Kaptoge S, Baldock P, Cooper C, Reeve J, Ntzani EE, Evangelou E, Ohlsson C, Karasik D, Rivadeneira F, Kiel DP, Tobias JH, Gregson CL, Harvey NC, Grundberg E, Goltzman D, Lelliott CJ, Hinds DA, Ackert-Bicknell CL, Hsu YH, Maurano MT, Croucher PI, Williams GR, Bassett JHD, Evans DM, Richards JB. An atlas of genetic influences on osteoporosis in humans and mice. *Nat Genet* 2019;**51**(2):258-266.

35. Wuttke M, Li Y, Li M, Sieber KB, Feitosa MF, Gorski M, Tin A, Wang L, Chu AY, Hoppmann A, Kirsten H, Giri A, Chai JF, Sveinbjornsson G, Tayo BO, Nutile T, Fuchsberger C, Marten J, Cocca M, Ghasemi S, Xu Y, Horn K, Noce D, van der Most PJ, Sedaghat S, Yu Z, Akiyama M, Afaq S, Ahluwalia TS, Almgren P, Amin N, Arnlov J, Bakker SJL, Bansal N, Baptista D, Bergmann S, Biggs ML, Biino G, Boehnke M, Boerwinkle E, Boissel M, Bottinger EP, Boutin TS, Brenner H, Brumat M, Burkhardt R, Butterworth AS, Campana E, Campbell A, Campbell H, Canouil M, Carroll RJ, Catamo E, Chambers JC, Chee ML, Chen X, Cheng CY, Cheng Y, Christensen K, Cifkova R, Ciullo M, Concas MP, Cook JP, Coresh J, Corre T, Sala CF, Cusi D, Danesh J, Daw EW, de Borst MH, De Grandi A, de Mutsert R, de Vries APJ, Degenhardt F, Delgado G, Demirkan A, Di Angelantonio E, Dittrich K, Divers J, Dorajoo R, Eckardt KU, Ehret G, Elliott P, Endlich K, Evans MK, Felix JF, Foo VHX, Franco OH, Franke A, Freedman BI, Freitag-Wolf S, Friedlander Y, Froguel P, Gansevoort RT, Gao H, Gasparini P, Gaziano JM, Giedraitis V, Gieger C, Girotto G, et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat Genet* 2019;**51**(6):957-972.

36. Paternoster L, Standl M, Waage J, Baurecht H, Hotze M, Strachan DP, Curtin JA, Bonnelykke K, Tian C, Takahashi A, Esparza-Gordillo J, Alves AC, Thyssen JP, den Dekker HT, Ferreira MA, Altmaier E, Sleiman PM, Xiao FL, Gonzalez JR, Marenholz I, Kalb B, Yanes MP, Xu CJ, Carstensen L, Groen-Blokhuis MM, Venturini C, Pennell CE, Barton SJ, Levin AM, Curjuric I, Bustamante M, Kreiner-Moller E, Lockett GA, Bacelis J, Bunyavanich S, Myers RA, Matanovic A, Kumar A, Tung JY, Hirota T, Kubo M, McArdle WL, Henderson AJ, Kemp JP, Zheng J, Smith GD, Ruschendorf F, Bauerfeind A, Lee-Kirsch MA, Arnold A, Homuth G, Schmidt CO, Mangold E, Cichon S, Keil T, Rodriguez E, Peters A, Franke A, Lieb W, Novak N, Folster-Holst R, Horikoshi M, Pekkanen J, Sebert S, Husemoen LL, Grarup N, de Jongste JC, Rivadeneira F, Hofman A, Jaddoe VW, Pasmans SG, Elbert NJ, Uitterlinden AG, Marks GB, Thompson PJ, Matheson MC, Robertson CF, Ried JS, Li J, Zuo XB, Zheng XD, Yin XY, Sun LD, McAleer MA, O'Regan GM, Fahy CM, Campbell LE, Macek M, Kurek M, Hu D, Eng C, Postma DS, Feenstra B, Geller F, Hottenga JJ, Middeldorp CM, Hysi P, Bataille V, Spector T, Tiesler CM, et al. Multi-ancestry genome-wide association study of 21,000 cases and

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

95,000 controls identifies new risk loci for atopic dermatitis. *Nat Genet* 2015;**47**(12):1449-1456.

37. Pulit SL, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, Yengo L, Ferreira T, Marouli E, Ji Y, Yang J, Jones S, Beaumont R, Croteau-Chonka DC, Winkler TW, Consortium G, Hattersley AT, Loos RJF, Hirschhorn JN, Visscher PM, Frayling TM, Yaghootkar H, Lindgren CM. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum Mol Genet* 2019;**28**(1):166-174.
38. Neale lab. http://www.nealelab.is.
39. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, Datta G, Davila-Velderrain J, McGuire D, Tian C, Zhan X, Choquet H, Docherty AR, Faul JD, Foerster JR, Fritsche LG, Gabrielsen ME, Gordon SD, Haessler J, Hottenga JJ, Huang H, Jang SK, Jansen PR, Ling Y, Magi R, Matoba N, McMahon G, Mulas A, Orru V, Palviainen T, Pandit A, Reginsson GW, Skogholt AH, Smith JA, Taylor AE, Turman C, Willemsen G, Young H, Young KA, Zajac GJM, Zhao W, Zhou W, Bjornsdottir G, Boardman JD, Boehnke M, Boomsma DI, Chen C, Cucca F, Davies GE, Eaton CB, Ehringer MA, Esko T, Fiorillo E, Gillespie NA, Gudbjartsson DF, Haller T, Harris KM, Heath AC, Hewitt JK, Hickie IB, Hokanson JE, Hopfer CJ, Hunter DJ, Iacono WG, Johnson EO, Kamatani Y, Kardia SLR, Keller MC, Kellis M, Kooperberg C, Kraft P, Krauter KS, Laakso M, Lind PA, Loukola A, Lutz SM, Madden PAF, Martin NG, McGue M, McQueen MB, Medland SE, Metspalu A, Mohlke KL, Nielsen JB, Okada Y, Peters U, Polderman TJC, Posthuma D, Reiner AP, Rice JP, Rimm E, Rose RJ, Runarsdottir V, Stallings MC, Stancakova A, Stefansson H, Thai KK, Tindle HA, Tyrfingsson T, Wall TL, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet* 2019;**51**(2):237-244.
40. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, Nguyen-Viet TA, Bowers P, Sidorenko J, Karlsson Linner R, Fontana MA, Kundu T, Lee C, Li H, Li R, Royer R, Timshel PN, Walters RK, Willoughby EA, Yengo L, Alver M, Bao Y, Clark DW, Day FR, Furlotte NA, Joshi PK, Kemper KE, Kleinman A, Langenberg C, Magi R, Trampush JW, Verma SS, Wu Y, Lam M, Zhao JH, Zheng Z, Boardman JD, Campbell H, Freese J, Harris KM, Hayward C, Herd P, Kumari M, Lencz T, Luan J, Malhotra AK, Metspalu A, Milani L, Ong KK, Perry JRB, Porteous DJ, Ritchie MD, Smart MC, Smith BH, Tung JY, Wareham NJ, Wilson JF, Beauchamp JP, Conley DC, Esko T, Lehrer SF, Magnusson PKE, Oskarsson S, Pers TH, Robinson MR, Thom K, Watson C, Chabris CF, Meyer MN, Laibson DI, Yang J, Johannesson M, Koellinger PD, Turley P, Visscher PM, Benjamin DJ, Cesarini D. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet* 2018;**50**(8):1112-1121.
41. Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, de Leeuw CA, Nagel M, Awasthi S, Barr PB, Coleman JRI, Grasby KL, Hammerschlag AR, Kaminski JA, Karlsson R, Krapohl E, Lam M, Nygaard M, Reynolds CA, Trampush JW, Young H, Zabaneh D, Hagg S, Hansell NK, Karlsson IK, Linnarsson S, Montgomery GW, Munoz-Manchado AB, Quinlan EB, Schumann G, Skene NG, Webb BT, White T, Arking DE, Avramopoulos D, Bilder RM, Bitsios P, Burdick KE, Cannon TD, Chiba-Falek O, Christoforou A, Cirulli ET, Congdon E, Corvin A, Davies G, Deary IJ, DeRosse P, Dickinson D, Djurovic S, Donohoe G, Conley ED, Eriksson JG, Espeseth T, Freimer NA, Giakoumaki S, Giegling I, Gill M, Glahn DC, Hariri AR, Hatzimanolis A, Keller MC,

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Knowles E, Koltai D, Konte B, Lahti J, Le Hellard S, Lencz T, Liewald DC, London E, Lundervold AJ, Malhotra AK, Melle I, Morris D, Need AC, Ollier W, Palotie A, Payton A, Pendleton N, Poldrack RA, Raikkonen K, Reinvang I, Roussos P, Rujescu D, Sabb FW, Scult MA, Smeland OB, Smyrnis N, Starr JM, Steen VM, Stefanis NC, Straub RE, Sundet K, Tiemeier H, Voineskos AN, Weinberger DR, Widen E, Yu J, Abecasis G, Andreassen OA, Breen G, Christiansen L, et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat Genet* 2018;**50**(7):912-919.

42. Chang-Quan H, Zheng-Rong W, Yong-Hong L, Yi-Zhou X, Qing-Xiu L. Education and risk for late life depression: a meta-analysis of published literature. *Int J Psychiatry Med* 2010;**40**(1):109-24.
43. Li Y, Cao J. Factors associated with suicidal behaviors in mainland China: a meta-analysis. *BMC Public Health* 2012;**12**:524.
44. Tang B, Deng Q, Glik D, Dong J, Zhang L. A Meta-Analysis of Risk Factors for Post-Traumatic Stress Disorder (PTSD) in Adults and Children after Earthquakes. *Int J Environ Res Public Health* 2017;**14**(12).
45. Wang MD, Little J, Gomes J, Cashman NR, Krewski D. Identification of risk factors associated with onset and progression of amyotrophic lateral sclerosis using systematic review and meta-analysis. *Neurotoxicology* 2017;**61**:101-130.
46. Xu W, Tan L, Wang HF, Jiang T, Tan MS, Zhao QF, Li JQ, Wang J, Yu JT. Meta-analysis of modifiable risk factors for Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 2015;**86**(12):1299-306.
47. Khaing W, Vallibhakara SA, Attia J, McEvoy M, Thakkinstian A. Effects of education and income on cardiovascular outcomes: A systematic review and meta-analysis. *Eur J Prev Cardiol* 2017;**24**(10):1032-1042.
48. Potter EL, Hopper I, Sen J, Salim A, Marwick TH. Impact of socioeconomic status on incident heart failure and left ventricular dysfunction: systematic review and meta-analysis. *Eur Heart J Qual Care Clin Outcomes* 2019;**5**(2):169-179.
49. McHutchison CA, Backhouse EV, Cvoro V, Shenkin SD, Wardlaw JM. Education, Socioeconomic Status, and Intelligence in Childhood and Stroke Risk in Later Life: A Meta-analysis. *Epidemiology* 2017;**28**(4):608-618.
50. Dong JY, Qin LQ. Education level and breast cancer incidence: a meta-analysis of cohort studies. *Menopause* 2019.
51. Brown CR, Hambleton I, Hercules SM, Unwin N, Murphy MM, Nigel Harris E, Wilks R, MacLeish M, Sullivan L, Sobers-Grannum N. Social determinants of prostate cancer in the Caribbean: a systematic review and meta-analysis. *BMC Public Health* 2018;**18**(1):900.
52. Sidorchuk A, Agardh EE, Aremu O, Hallqvist J, Allebeck P, Moradi T. Socioeconomic differences in lung cancer incidence: a systematic review and meta-analysis. *Cancer Causes Control* 2009;**20**(4):459-71.
53. Bellou V, Belbasis L, Tzoulaki I, Evangelou E. Risk factors for type 2 diabetes mellitus: An exposure-wide umbrella review of meta-analyses. *PLoS One* 2018;**13**(3):e0194127.
54. Zeng X, Liu J, Tao S, Hong HG, Li Y, Fu P. Associations between socioeconomic status and chronic kidney disease: a meta-analysis. *J Epidemiol Community Health* 2018;**72**(4):270-279.
55. He J, Chen X, Fan X, Cai Z, Huang F. Is there a relationship between body mass index and academic achievement? A meta-analysis. *Public Health* 2019;**167**:111-124.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

56. Leng B, Jin Y, Li G, Chen L, Jin N. Socioeconomic status and hypertension: a meta-analysis. *J Hypertens* 2015;**33**(2):221-9.
57. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* 2016;**40**(4):304-14.
58. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol* 2017;**32**(5):377-389.
59. Davies NM, Hill WD, Anderson EL, Sanderson E, Deary IJ, Davey Smith G. Multivariable two-sample Mendelian randomization estimates of the effects of intelligence and education on health. *Elife* 2019;**8**.
60. Rees JMB, Wood AM, Burgess S. Extending the MR-Egger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy. *Stat Med* 2017;**36**(29):4705-4718.
61. Carter AR, Sanderson E, Hammerton G, Richmond RC, Smith GD, Heron J, Taylor AE, Davies NM, Howe LD. Mendelian randomisation for mediation analysis: current methods and challenges for implementation. 2019.
62. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol* 2013;**42**(5):1497-501.
63. Spiller W, Davies NM, Palmer TM. Software Application Profile: mrrobust - a tool for performing two-sample summary Mendelian randomization analyses. 2017.
64. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J, Langdon R, Tan VY, Yarmolinsky J, Shihab HA, Timpson NJ, Evans DM, Relton C, Martin RM, Davey Smith G, Gaunt TR, Haycock PC. The MR-Base platform supports systematic causal inference across the human phenome. *Elife* 2018;**7**.
65. Sterne JA, Davey Smith G. Sifting the evidence-what's wrong with significance tests? *Bmj* 2001;**322**(7280):226-31.
66. Lund Nilsen TI, Johnsen R, Vatten LJ. Socio-economic and lifestyle factors associated with the risk of prostate cancer. *Br J Cancer* 2000;**82**(7):1358-63.
67. Zhou H, Zhang Y, Liu J, Yang Y, Fang W, Hong S, Chen G, Zhao S, Zhang Z, Shen J, Xian W, Huang Y, Zhao H, Zhang L. Education and lung cancer: a Mendelian randomization study. *Int J Epidemiol* 2019;**48**(3):743-750.
68. Tillmann T, Vaucher J, Okbay A, Pikhart H, Peasey A, Kubinova R, Pajak A, Tamosiunas A, Malyutina S, Hartwig FP, Fischer K, Veronesi G, Palmer T, Bowden J, Davey Smith G, Bobak M, Holmes MV. Education and coronary heart disease: mendelian randomisation study. *Bmj* 2017;**358**:j3542.
69. Larsson SC, Traylor M, Malik R, Dichgans M, Burgess S, Markus HS. Modifiable pathways in Alzheimer's disease: Mendelian randomisation analysis. *Bmj* 2017;**359**:j5375.
70. Bockerman P, Viinikainen J, Pulkki-Raback L, Hakulinen C, Pitkanen N, Lehtimaki T, Pehkonen J, Raitakari OT. Does higher education protect against obesity? Evidence using Mendelian randomization. *Prev Med* 2017;**101**:195-198.
71. Gage SH, Bowden J, Davey Smith G, Munafo MR. Investigating causality in associations between education and smoking: a two-sample Mendelian randomization study. *Int J Epidemiol* 2018;**47**(4):1131-1140.
72. Carter AR, Gill D, Davies NM, Taylor AE, Tillmann T, Vaucher J, Wootton RE, Munafo MR, Hemani G, Malik R, Seshadri S, Woo D, Burgess S, Davey Smith G, Holmes MV, Tzoulaki I, Howe LD, Dehghan A. Understanding the consequences of education

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

inequality on cardiovascular disease: mendelian randomisation study. *Bmj* 2019;**365**:l1855.

73. Kowalec K, Lu Y, Sariaslan A, Song J, Ploner A, Dalman C, Hultman CM, Larsson H, Lichtenstein P, Sullivan PF. Increased schizophrenia family history burden and reduced premorbid IQ in treatment-resistant schizophrenia: a Swedish National Register and Genomic Study. *Mol Psychiatry* 2019.
74. Gale CR, Batty GD, McIntosh AM, Porteous DJ, Deary IJ, Rasmussen F. Is bipolar disorder more common in highly intelligent people? A cohort study of a million men. *Mol Psychiatry* 2013;**18**(2):190-4.
75. Smeland OB, Bahrami S, Frei O, Shadrin A, O'Connell K, Savage J, Watanabe K, Krull F, Bettella F, Steen NE, Ueland T, Posthuma D, Djurovic S, Dale AM, Andreassen OA. Genome-wide analysis reveals extensive genetic overlap between schizophrenia, bipolar disorder, and intelligence. *Mol Psychiatry* 2019.
76. Zeng L, Ntalla I, Kessler T, Kastrati A, Erdmann J, Danesh J, Watkins H, Samani NJ, Deloukas P, Schunkert H. Genetically modulated educational attainment and coronary disease risk. *Eur Heart J* 2019;**40**(29):2413-2420.
77. Stringhini S, Polidoro S, Sacerdote C, Kelly RS, van Veldhoven K, Agnoli C, Grioni S, Tumino R, Giurdanella MC, Panico S, Mattiello A, Palli D, Masala G, Gallo V, Castagne R, Paccaud F, Campanella G, Chadeau-Hyam M, Vineis P. Life-course socioeconomic status and DNA methylation of genes regulating inflammation. *Int J Epidemiol* 2015;**44**(4):1320-30.
78. Huang JY, Gavin AR, Richardson TS, Rowhani-Rahbar A, Siscovick DS, Hochner H, Friedlander Y, Enquobahrie DA. Accounting for Life-Course Exposures in Epigenetic Biomarker Association Studies: Early Life Socioeconomic Position, Candidate Gene DNA Methylation, and Adult Cardiometabolic Risk. *Am J Epidemiol* 2016;**184**(7):520-531.
79. Swartz JR, Hariri AR, Williamson DE. An epigenetic mechanism links socioeconomic status to changes in depression-related brain function in high-risk adolescents. *Mol Psychiatry* 2017;**22**(2):209-214.
80. Song L, Meng J, Liu Q, Huo T, Zhu X, Li Y, Ren Z, Wang X, Qiu J. Polygenic Score of Subjective Well-Being Is Associated with the Brain Morphology in Superior Temporal Gyrus and Insula. *Neuroscience* 2019;**414**:210-218.
81. Okbay A, Baselmans BM, De Neve JE, Turley P, Nivard MG, Fontana MA, Meddens SF, Linner RK, Rietveld CA, Derringer J, Gratten J, Lee JJ, Liu JZ, de Vlaming R, Ahluwalia TS, Buchwald J, Cavadino A, Frazier-Wood AC, Furlotte NA, Garfield V, Geisel MH, Gonzalez JR, Haitjema S, Karlsson R, van der Laan SW, Ladwig KH, Lahti J, van der Lee SJ, Lind PA, Liu T, Matteson L, Mihailov E, Miller MB, Minica CC, Nolte IM, Mook-Kanamori D, van der Most PJ, Oldmeadow C, Qian Y, Raitakari O, Rawal R, Realo A, Rueedi R, Schmidt B, Smith AV, Stergiakouli E, Tanaka T, Taylor K, Thorleifsson G, Wedenoja J, Wellmann J, Westra HJ, Willems SM, Zhao W, Amin N, Bakshi A, Bergmann S, Bjornsdottir G, Boyle PA, Cherney S, Cox SR, Davies G, Davis OS, Ding J, Direk N, Eibich P, Emeny RT, Fatemifar G, Faul JD, Ferrucci L, Forstner AJ, Gieger C, Gupta R, Harris TB, Harris JM, Holliday EG, Hottenga JJ, De Jager PL, Kaakinen MA, Kajantie E, Karhunen V, Kolcic I, Kumari M, Launer LJ, Franke L, Li-Gao R, Liewald DC, Koini M, Loukola A, Marques-Vidal P, Montgomery GW, Mosing MA, Paternoster L, Pattie A, Petrovic KE, Pulkki-Raback L, Quaye L, Raikkonen K, Rudan I, Scott RJ, et al.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.01.20020008; this version posted September 13, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat Genet* 2016;**48**(6):624-33.

82. Baselmans BML, Bartels M. A genetic perspective on the relationship between eudaimonic -and hedonic well-being. *Sci Rep* 2018;**8**(1):14610.
83. Day FR, Ong KK, Perry JRB. Elucidating the genetic basis of social interaction and isolation. *Nat Commun* 2018;**9**(1):2457.
84. Richardson JG. *Handbook of theory and research for the sociology of education*. Westport, Conn.: Greenwood Press, 1986.
85. Bourdieu P. *Distinction : a social critique of the judgement of taste*. London: Routledge, 2010.
86. Manrique-Garcia E, Sidorchuk A, Hallqvist J, Moradi T. Socioeconomic position and incidence of acute myocardial infarction: a meta-analysis. *J Epidemiol Community Health* 2011;**65**(4):301-9.
87. Stringhini S, Sabia S, Shipley M, Brunner E, Nabi H, Kivimaki M, Singh-Manoux A. Association of socioeconomic position with health behaviors and mortality. *Jama* 2010;**303**(12):1159-66.