**Figure legends**

**Figure 1.  Conceptual framework of the study: Illustration of identifying novel factors that influence the original association.** (a) Where (gy) is the total effect of the SNP on the outcome, (gx) is the SNP-exposure effect, (xy) is the exposure-outcome effect as estimated through MR analysis from the non-outlier SNPs, (gp) is the SNP-candidate trait effect and (py) is the causal effect of the candidate trait on the outcome. (b) The open circles represent valid instruments and the slope of the dotted line represents the causal effect estimate of the exposure on the outcome. The closed circle represents an outlier SNP which influences the outcome, through two independent pathways (py).(c) One way in which the red SNP can exhibit a larger influence on the outcome than expected given its effect on the exposure is if it influences the outcome additionally through another pathway (horizontal pleiotropy). (d) Using the MR-Base database of GWAS summary data for hundreds of traits we can search for ‘candidate traits’ with which the outlier SNP has an association. (e) The causal inference of each of those candidate traits on the outcome can be estimated using MR by identifying their instruments (excluding the original outlier SNP). This allows us to identify new traits that putatively influence the outcome.

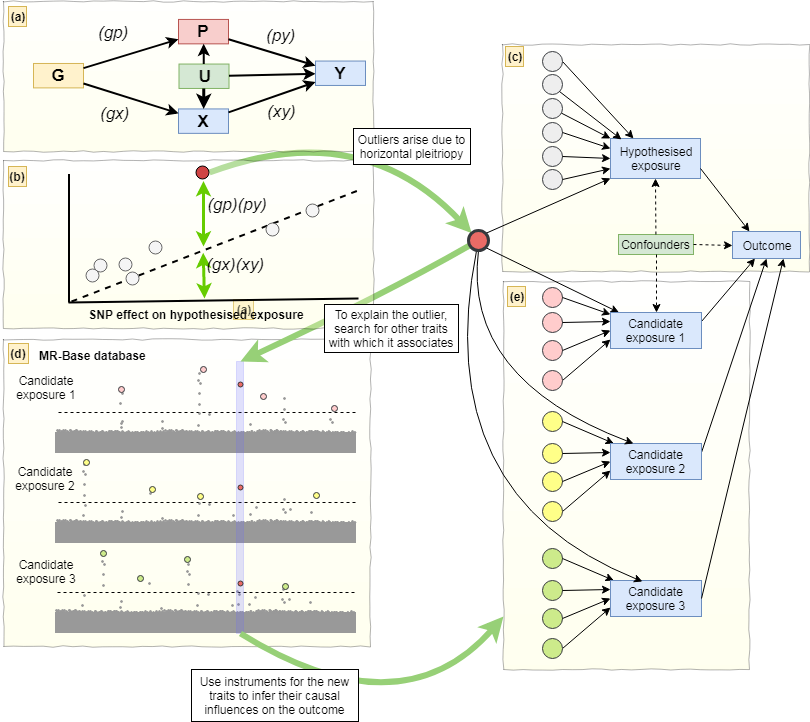
**Figure 2. Result from simulations.** (A) False discovery rates (first panel where the simulated causal effect is 0) and power (second panel where the simulated causal effect is 0.2) to detect an association between simulated exposure and outcome traits. The x-axis depicts the number of instruments (out of 30) that exhibit a horizontal pleiotropic effect. There are four methods for handling outliers, and we posit two scenarios for detecting outliers. The methods are ‘raw’, where all SNPs are used in a standard IVW analysis regardless of outlier status; ‘outliers adjusted’ where the outlier SNPs are adjusted for detected alternative pathways; ‘outliers removed (all)’ where all detected outliers are removed; and ‘outliers removed (candidate)’ where only outliers that are found to influence a candidate trait are removed. We run the latter three methods by detecting outliers empirically, but also show, for comparison, the hypothetical case in which we know the pleiotropic variants *a priori*. (B) As in (A), except comparing the bias of different methods, assessed as the proportion of estimates that are substantially different from the simulated effect (y-axis).

**Figure 3. The results of empirical analyses.** A scatter plot relating the effect size of each SNP-exposure association and the SNP-outcome association with 95% confidence intervals. The slopes of the line correspond to causal estimates using each of two different methods. (A) Empirical analysis 1: Systolic blood pressure (mmHg) and coronary heart disease (log odds). (B) Empirical analysis 2: Years of schooling (years) and body mass index (kg/m2). (C) Empirical analysis 3: Urate (mg/dl) and coronary heart disease (log odds). (D) Empirical analysis 4: Sleep duration (hour/night) and schizophrenia (log odds).

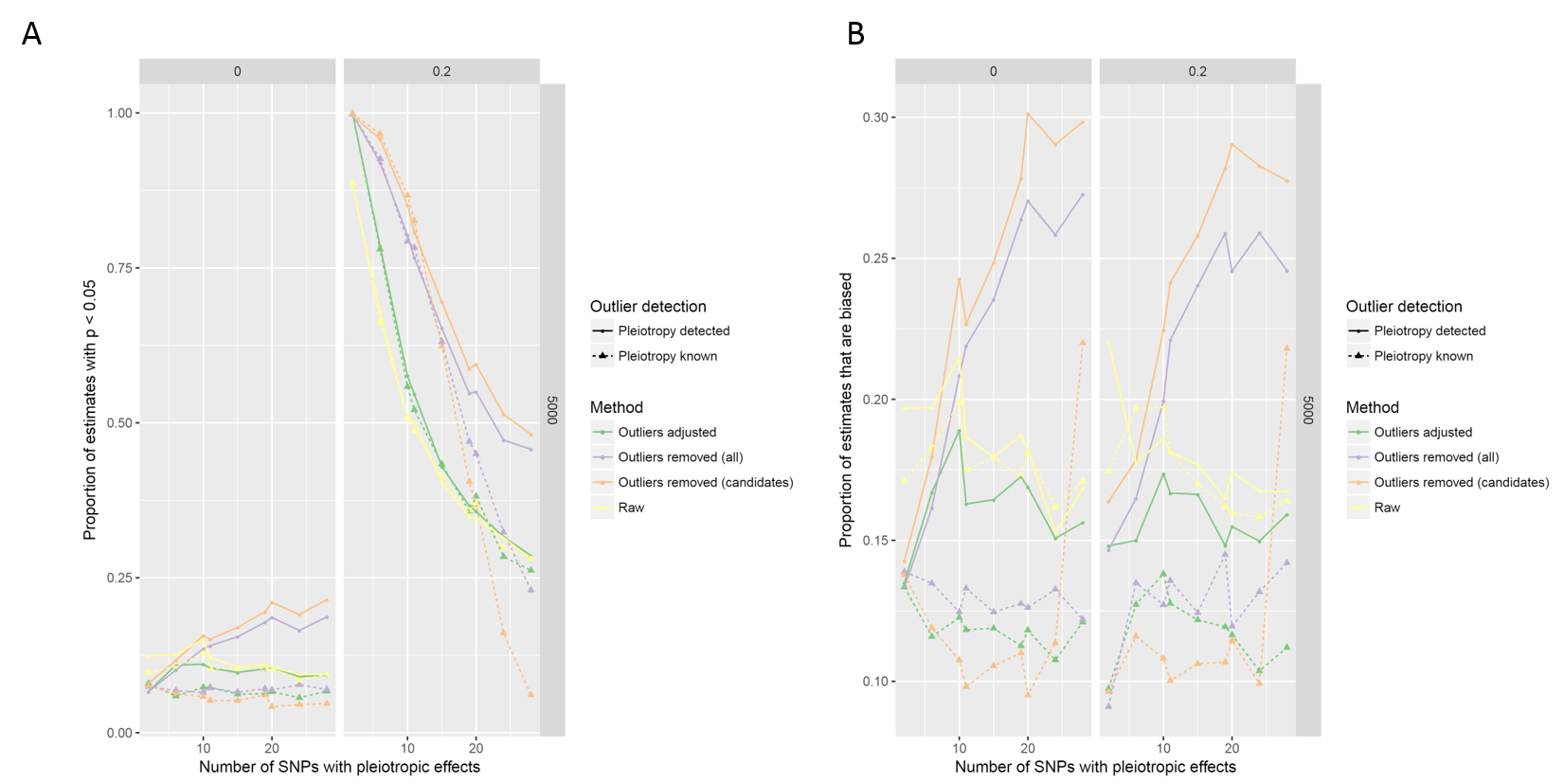
**Figure 4. Manhattan plot to visualise the causal associations between candidate exposures and (A) hypothesised exposure or (B) hypothesised outcome.** This represents the number of traits associated with outliers. The plot is stratified by phenotype category and, within each group, we present the results related to the candidate traits identified. Along the X axis, different phenotype groups are shown in different colours. The Y axis presents log transformed P value for each trait. Filled circles in each category indicate the evidence of association between candidate traits and exposure or outcome (p < 0.05). (A) Empirical analysis 1: Systolic blood pressure (mmHg) and coronary heart disease (log odds). (B) Empirical analysis 2: Years of schooling (years) and body mass index (kg/m2). (C) Empirical analysis 3: Urate (mg/dl) and coronary heart disease (log odds). (D) Empirical analysis 4: Sleep duration (hour/night) and schizophrenia (log odds).

**Figure 5. Scatter plot for the exposure-outcome association adjusting the SNP effects on the candidate traits.** The arrow indicates changes in the SNP effect after conditioning on the effect of candidate traits on the outcome. The candidate traits that influence the association of the original exposure and the original outcome were listed in the box. (A) Empirical analysis 1: Systolic blood pressure (mmHg) and coronary heart disease (log odds). (B) Empirical analysis 2: Years of schooling (years) and body mass index (kg/m2). (C) Empirical analysis 3: Urate (mg/dl) and coronary heart disease (log odds). (D) Empirical analysis 4: Sleep duration (hour/night) and schizophrenia (log odds).

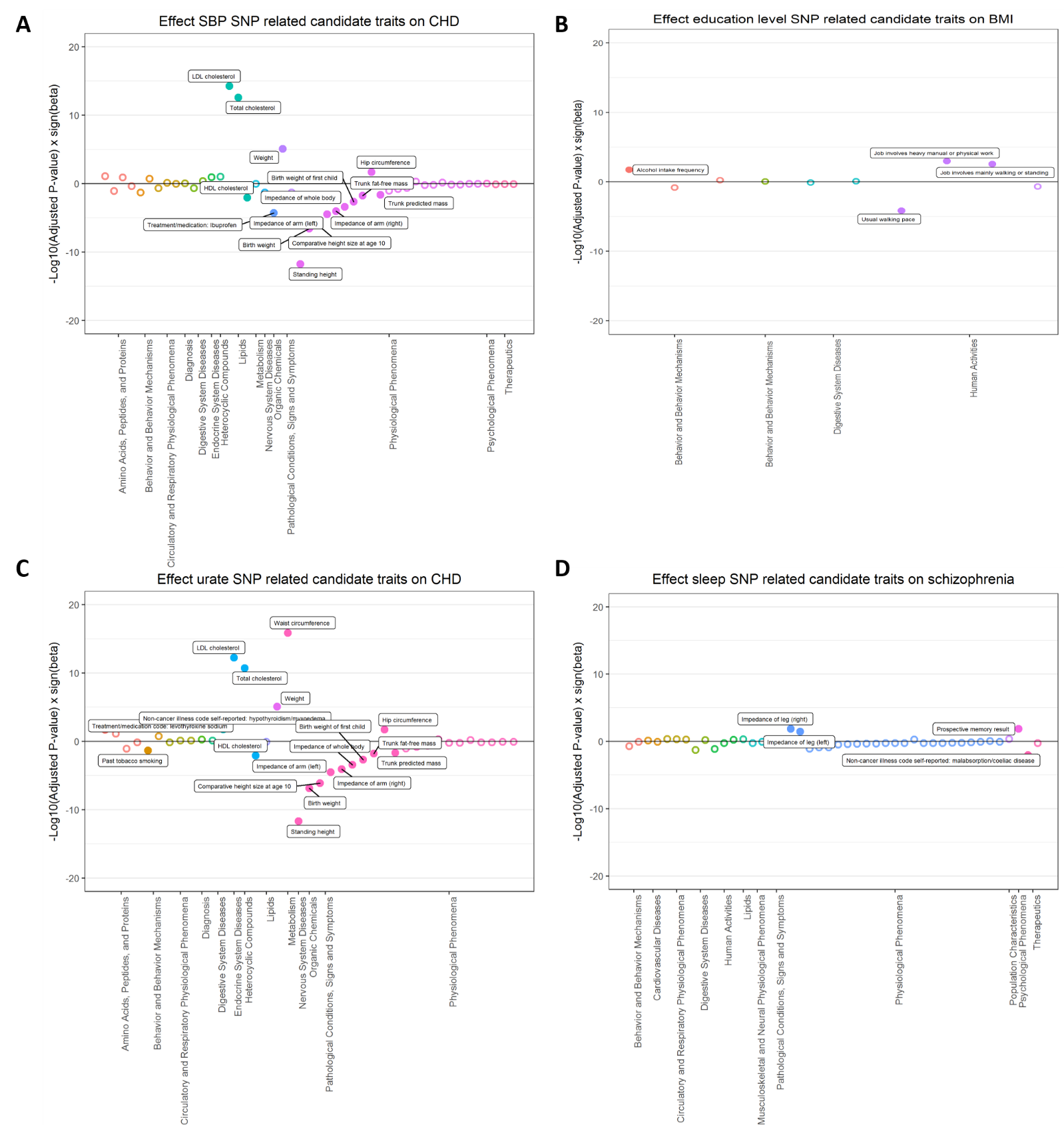
**Figure 1.**



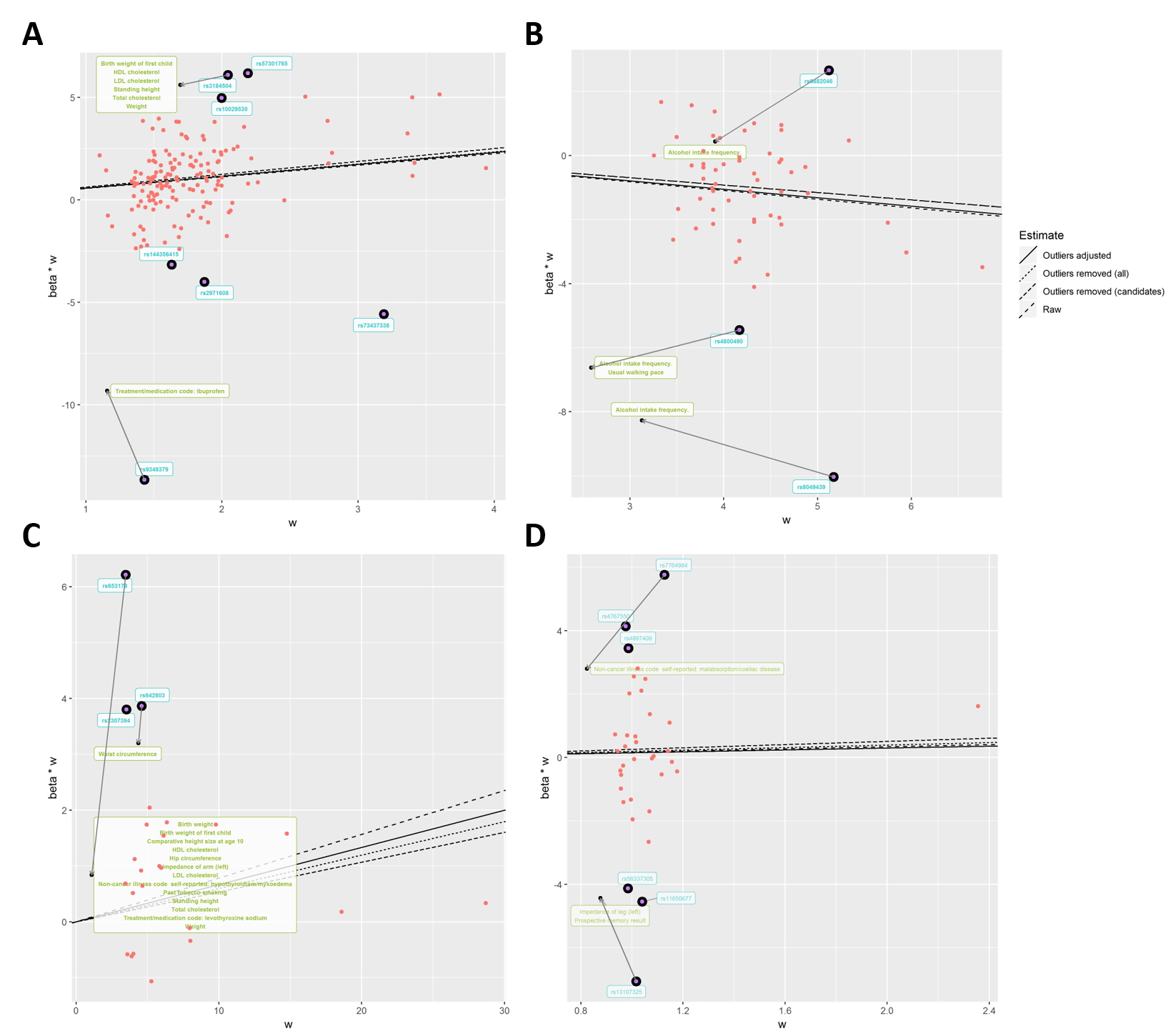
**Figure 2.**



**Figure 3.**



**Figure 4.**



**Table 1. Candidate traits associated with both exposure and outcome.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Outlier SNPs** | **Nearest gene** | **Category** | **Phenotypes** 1 | **N SNPs** 2 | **Beta (95% CI)** 3 |
| **Empirical analysis 1. Systolic blood pressure (mmHg) and coronary heart disease (log odds)** | | | | | |
| rs3184504 | *SH2B3* | Early development | Birth weight of first child | 40 | -0.312 (-0.498, -0.126) |
| Anthropometric measures | Standing height | 577 | -0.208 (-0.264, -0.152) |
| Lipid | LDL cholesterol  HDL cholesterol  Total cholesterol | 78  86  86 | 0.393 (0.290, 0.497)  -0.172 (-0.288, -0.055)  0.378 (0.271, 0.484) |
| rs9349379 | *PHACTR* | Medications | Ibuprofen | 1 | -7.790 (-11.340, -4.239) |
| **Empirical analysis 2. Years of schooling (years) and body mass index (kg/m2)** | | | | | |
| rs6882046 | *LINC00461* | Drinking | Alcohol intake frequency | 31 | 0.347 (0.065, 0.628) |
| rs4800490 | *NPC1* | Drinking | Alcohol intake frequency | 31 | 0.347 (0.065, 0.628) |
| Exercise | Usual walking pace | 22 | -1.595 (-2.364, -0.825) |
| rs8049439 | *ATXN2L* | Drinking | Alcohol intake frequency | 31 | 0.347 (0.065, 0.628) |
| **Empirical analysis 3. Urate (mg/dl) and coronary heart disease (log odds)** | | | | | |
| rs653178 | *ATXN2* | Early development | Birth weight of first child  Birth weight | 31  40 | 0.347 (0.065, 0.628)  -0.312 (-0.498, -0.126) |
| Anthropometric measures | Comparative height size at age 10  Hip circumference  Impedance of arm (left)  Standing height | 357  275  305  577 | -0.248 (-0.342, -0.154)  0.131 (0.030, 0.231)  -0.263 (-0.38, -0.145)  -0.208 (-0.264, -0.152) |
| Lipid | HDL cholesterol  LDL cholesterol  Total cholesterol | 78  86  86 | 0.393 (0.290, 0.497)  -0.172 (-0.288, -0.055)  0.378 (0.271, 0.484) |
| Disease | hypothyroidism/myxoedema (Self-reported) | 77 | 0.847 (0.211, 1.483) |
| Smoking | Past tobacco smoking | 41 | -0.265 (-0.5, -0.029) |
| Medications | Treatment/medication: levothyroxine sodium | 51 | 1.231 (0.27, 2.191) |
| rs642803 | *OVOL1* | Anthropometric measures | Waist circumference | 218 | 0.458 (0.352, 0.563) |
| **Empirical analysis 4. Sleep duration (hour/night) and schizophrenia (log odds)** | | | | | |
| rs7764984 | *HIST1H2BJ* | Disease | Malabsorption/coeliac disease (self-reported) | 11 | -8.401 (-12.842, -3.961) |
| rs13107325 | *SLC39A8* | Anthropometric measures | Impedance of leg (left) | 282 | 0.179 (0.047, 0.311) |
| Memory | Prospective memory result | 2 | 4.493 (1.851, 7.135) |

SNP, single nucleotide polymorphism; VLDL, very low-density lipoprotein; HDLC, high density lipoprotein cholesterol; LDLC, low density lipoprotein cholesterol; N SNPs, number of SNPs; CI, confidence interval.

1 Candidate traits that are associated with outliers (p < 5 x 10-8) and both exposure and outcome are listed. The listed traits were used in the adjusted model to investigate whether they are associated with CHD. 2 The number of SNPs used for two sample MR analysis of candidate traits on the outcome. 3 The results were presented as IVW beta coefficient (95% CI), derived from two sample MR analyses. The estimate for self-reported ibuprofen use obtained from Wald ratio method since IVW method is not available when the number of instruments is less than 2.

**Table 2. The results of empirical analyses with different IV estimators derived from different methods.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | **Estimate (95% CIs)** | | | |
| **Method** | | **All variants** | **Removing outliers** | **Removing candidate outliers** | **Adjustment for candidate outliers** |
| **Empirical analysis 1. Systolic blood pressure (mmHg) and coronary heart disease (log odds)** | | | | | |
|  | Heterogeneity (*Q*)1 | 682.7  (N SNPs = 157) | 312.1  (N SNPs = 150) | 448.7  (N SNPs = 155) | 567.6  (N SNPs = 157) |
|  | IVW random effects | 0.566 (0.388, 0.744) | 0.629 (0.504, 0.754) | 0.586 (-0.584, 1.756) | 0.534 (0.371, 0.697) |
|  | Egger random effects | 0.971 (0.399, 1.543) | 1.082 (0.678, 1.486) | 0.791 (-1.157, 2.739) | - |
|  | Intercept | -0.020 (-0.032, -0.008) | -0.010 (-0.018, -0.002) | -0.004 (-0.012, 0.004) | - |
|  | Weighted median | 0.571 (0.424, 0.718) | 0.578 (0.431, 0.725) | 0.568 (-0.551, 1.687) | - |
|  | Weighted mode | 0.571 (0.234, 0.908) | 0.546 (0.197, 0.895) | 0.554 (-0.510, 1.618) | - |
| **Empirical analysis 2. Years of schooling (years) and body mass index (kg/m2)** | | | | | |
|  | Heterogeneity (*Q*) | 211.9  (N SNPs = 59) | 101.9  (N SNPs = 56) | 101.9  (N SNPs = 56) | 197.8  (N SNPs =59) |
|  | IVW random effects | -0.272 (-0.386, -0.158) | -0.232 (-0.314, -0.150) | -0.232 (-0.314, -0.150) | -0.265 (-0.377, -0.153) |
|  | Egger random effects | 0.013 (-0.677, 0.703) | -0.404 (-0.910, 0.102) | -0.404 (-0.910, 0.102) | - |
|  | Intercept | -0.005 (-0.017, 0.007) | 0.003 (-0.005, 0.011) | 0.003 (-0.005, 0.011) | - |
|  | Weighted median | -0.209 (-0.307, -0.111) | -0.217 (-0.315, -0.119) | -0.217 (-0.315, -0.119) | - |
|  | Weighted mode | -0.141 (-0.413, 0.131) | -0.127 (-0.405, 0.151) | -0.127 (-0.405, 0.151) | - |
| **Empirical analysis 3. Urate (mg/dl) and coronary heart disease (log odds)** | | | | | |
|  | Heterogeneity (*Q*) | 81.6  (N SNPs = 24) | 20.7  (N SNPs = 21) | 33.4  (N SNPs = 22) | 44.1  (N SNPs =24) |
|  | IVW random effects | 0.078 (-0.004, 0.160) | 0.053 (0.008, 0.098) | 0.060 (0.055, 0.115) | 0.068 (-0.008, 0.144) |
|  | Egger random effects | -0.049 (-0.167, 0.069) | 0.008 (-0.065, 0.081) | -0.010 (-0.094, 0.074) | - |
|  | Intercept | 0.015 (0.003, 0.027) | 0.006 (-0.002, 0.014) | -0.008 (-0.016, 0.0002) | - |
|  | Weighted median | 0.019 (-0.040, 0.078) | 0.016 (-0.043, 0.075) | 0.017 (-0.040, 0.074) | - |
|  | Weighted mode | 0.028 (-0.025, 0.081) | 0.022 (-0.035, 0.079) | 0.025 (-0.030, 0.080) | - |
| **Empirical analysis 4. Sleep duration (hour/night) and schizophrenia (log odds)** | | | | | |
|  | Heterogeneity (*Q*) | 204.8  (N SNPs = 36) | 54.1  (N SNPs = 30) | 121.4  (N SNPs = 34) | 147.7  (N SNPs =36) |
|  | IVW random effects | 0.169 (-0.556, 0.894) | 0.254 (-0.189, 0.697) | 0.195 (-0.395, 0.785) | 0.166 (-0.455, 0.787) |
|  | Egger random effects | -0.144 (-2.882, 2.594) | 0.887 (-0.724, 2.498) | 0.860 (-1.369, 3.089) | - |
|  | Intercept | 0.004 (-0.033, 0.041) | -0.009 (-0.031, 0.013) | -0.009 (-0.038, 0.020) | - |
|  | Weighted median | 0.244 (-0.256, 0.744) | 0.222 (-0.293, 0.737) | 0.223 (-0.273, 0.719) | - |
|  | Weighted mode | 0.283 (-0.387, 0.953) | 0.408 (-0.317, 1.133) | 0.356 (-0.354, 1.066) | - |

N SNPs, number of single nucleotide polymorphisms; 95% CIs, 95% confidence intervals; IVW, Inverse variance weighted. 1 Heterogeneity amongst the estimates were assessed based on contribution of individual variant to Cochran’s statistic.