Appraising the causal role of risk factors in coronary artery disease and stroke

Appraising the causal role of risk factors in coronary artery disease and stroke: A systematic review of Mendelian Randomization studies

Andrea N Georgiou[1], MSc, PhD, Loukas Zagkos[2], MSc, PhD, Georgios Markozannes[1,2,] , MSc, PhD, Christos V. Chalitsios[1], MSc, PhD, Alexandros Georgios Asimakopoulos[1], MSc, Wei Xu[3], MSc, PhD,, Lijuan Wang[3], MSc, PhD, Ines Mesa-Eguiagaray[3], MSc, PhD, Xuan Zhou[3], MBBS, Eleni M Loizidou[1,5], MSc, PhD,, Nikolaos Kretsavos[1] MD, Evropi Theodoratou[3,4], MSc, PhD,, Dipender Gill[2,6], BMBCh, PhD, Stephen Burgess[6,7], MMath, PhD, Evangelos Evangelou[1,2,8], MSc, PhD, Konstantinos K Tsilidis[1,2], MSc, PhD, Ioanna Tzoulaki[2,9], MSc, PhD

1. Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece.
2. Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom.
3. Centre for Global Health, Usher Institute, The University of Edinburgh, Edinburgh, UK.
4. CRUK Edinburgh Centre, Institute of Genetics and Cancer, The University of Edinburgh, Edinburgh, UK
5. Biobank.cy Center of Excellence in Biobanking and Biomedical Research, University of Cyprus
6. Medical Research Council Biostatistics Unit, University of Cambridge, Cambridge, UK
7. Cardiovascular Epidemiology Unit, University of Cambridge, Cambridge, UK
8. Department of Biomedical Research, Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, Ioannina, Greece
9. Centre for Systems Biology, Biomedical Research Foundation, Academy of Athens, Athens, Greece

Corresponding author

Ioanna Tzoulaki

Centre for Systems Biology, Biomedical Research Foundation, Academy of Athens, Athens, Greece itzoulaki@bioacademy.gr

Abstract

Background: Mendelian randomization (MR) offers a powerful approach to study potential causal associations between exposures and health outcomes, by using genetic variants associated with an exposure as instrumental variables. In this systematic review, we aimed to summarize previous MR studies and to evaluate the evidence for causality for a broad range of exposures in relation to coronary artery disease (CAD) and stroke.

Methods and Results: MR studies investigating the association of any genetically predicted exposure with CAD or stroke were identified. Studies were classified into four categories, namely *robust*, *probable*, *suggestive* and *insufficient*, built on the significance of the main MR analysis results and its concordance with sensitivity analyses. Associations that did not perform any sensitivity analysis were classified as *non-evaluable*. We identified 2,725 associations eligible for evaluation, examining 535 distinct exposures. Of them, 141 were classified as *robust*, 353 as *probable*, 110 as *suggestive* and 926 had *insufficient* evidence. The most prominent *robust* associations were observed for anthropometric traits, lipids and lipoproteins and type 2 diabetes with CAD, clinical measurements with CAD and stroke, and thrombotic factors with stroke.

Conclusions: Despite the large number of studies that have been conducted, only a limited number of associations were supported by robust evidence. About half of the associations presented a MR sensitivity analysis along with the main analysis which further supported the causality of associations. Future research should focus on more thorough assessment of sensitivity MR analyses and further assessment of mediation effects or nonlinearity of associations.

Keywords

Mendelian randomization; systematic review; cardiovascular disease; evidence grading

Clinical Perspective

What is new?

- Numerous Mendelian Randomization studies have examined the potential causal associations between risk factors and coronary artery disease (CAD) or stroke; robust evidence for causality has been shown for only a minority of them.
- The findings of this systematic review suggest that coronary artery disease and stroke share a somewhat different profile of robust associations with protective/risk factors; CAD was robustly associated with anthropometric traits and lipoproteins, whereas stroke was robustly associated with inflammatory biomarkers and thrombotic factors.

What Are the Clinical Implications?

- Future research should comply with reporting guidelines for Mendelian Randomization studies and focus on more thorough assessment of sensitivity MR analyses as well as other methodologies including investigation of mediation effects and nonlinearity of associations.

- Risk factors with robust support for causality should be further investigated to guide better cardiovascular disease prevention policies and treatment.

Non-standard Abbreviations and Acronyms

BMI: body mass index

CAD: coronary artery disease

CI: confidence interval

CVD: cardiovascular disease

CARDIoGRAMplusC4D: Coronary ARtery DIsease Genome wide Replication and Meta-analysis (CARDIoGRAM) plus The Coronary Artery Disease (C4D) Genetics consortium

CRP: C-reactive protein

GWAS: genome-wide association study

ICA1L: Islet cell autoantigen 1-like protein

IL6R: interleukin 6 receptor

IVW: inverse variance weighted

MR: mendelian randomization

MVMR: multivariate mendelian randomization

NK3R: neurokinin 3 receptor

RCTs: randomized controlled trials

sIL6R: soluble IL6 receptor

STROBE-MR: Strengthening the Reporting of Mendelian Randomization Studies

WM: weighted median

Introduction

Cardiovascular disease (CVD), principally coronary artery disease (CAD) and stroke, is the leading cause of death globally and a major contributor to disability worldwide [1]. A large body of research has concentrated on identifying risk factors for CAD and stroke since the early cardiovascular observational studies in 1950s [2]. These studies were instrumental in establishing the so-called conventional cardiovascular risk factors such as raised blood pressure, raised serum cholesterol, cigarette smoking and diabetes mellitus. However, beyond these conventional risk factors, an ever-expanding list of exposures and their associations with cardiovascular manifestations is being explored in the medical literature.

Despite the volume of research, the causality of associations between risk factors and cardiovascular outcomes remains inconclusive for the majority of exposures, as observational associations are hindered by confounding and reverse causation and evidence from randomized controlled trials (RCTs) is relatively scarce [3]. The Mendelian randomization (MR) approach can potentially overcome some biases of traditional epidemiological research by using genetic variants robustly associated with the risk factor of interest and assessing whether these variants are associated with the outcome of interest. The MR method can address bias due to confounding because genetic variants are randomly allocated when alleles are passed from parents to offspring during meiosis. MR studies therefore can be thought as 'randomizing' participants based on the presence of alleles which influence the risk factors of interest and subsequently investigate whether carriers of genetic variants associated with the risk factor have different disease risk compared to non-carriers. Additionally, as genetic variants are acquired at birth and cannot be modified by the presence of disease, MR associations are not influenced by reverse causality. Due to these appealing properties and as genome-wide association studies (GWAS) provide associations between numerous traits and risk factors, MR

is increasingly becoming a popular method to study the potential causal associations between different exposures and cardiovascular outcomes.

In this study we present the first effort to systematically collect and appraise MR studies investigating any risk factor in relation to CAD and stroke. Our aim was to present the breadth and depth of exposures studied, identify areas of research focus and highlight gaps, and to appraise the current evidence supporting their causal role in developing CAD and stroke.

Methods

The data and materials that support the findings of this study are available in the Supplementary material. Institutional review committee approval and consent by participants, was not required because the current study is based exclusively on summary level data from previous published studies.

Search Strategy:

A systematic literature search was conducted independently by two researchers (A.G and N.K) on Medline (via PubMed) from inception up to May 2022, for the identification of studies using the MR approach investigating causal risk factors for CAD and/or stroke. The following algorithm was used: "(Mendelian Randomization OR Mendelian Randomisation or genetic instrument) AND (Cardiovascular OR Stroke OR Coronary Heart OR Coronary Artery OR Myocardial Infarction)". We also screened the references of relevant reviews and the references of the included studies. The screening process in shown in Figure 1. The inclusion and exclusion criteria are described in detail in the Supplemental Material - Data S1.

Data Extraction

Data extraction was performed independently by four investigators (AG, LZ, CVC, AGA, EL, LW), and independently double-checked by two additional investigators (WX, IME). From each eligible article, we recorded the first author, year of publication, the examined risk factors (exposure) and corresponding outcomes, the sample size for each exposure and outcome, the exposure and outcome population ancestry and source (i.e. name of consortium), information for genetic variants modelled as instruments (p-value threshold, threshold for linkage disequilibrium, biological relevance, power, percentage of variance explained by the instruments), MR design, main MR analysis used (i.e. inverse variance weighted (IVW), maximum likelihood, Wald ratio, two-stage least squares), the effect size (odds ratio) and the corresponding 95% confidence interval (CI) and p-value. We further extracted information on a number of sensitivity MR methods, whenever these were performed and reported e.g., MR-Egger, weighted median (WM), MR-PRESSO and also multivariable MR (MVMR) and all items included in the Strengthening the Reporting of Mendelian Randomization Studies (STROBE-MR) checklist [4,5]. Outcomes included CAD (including myocardial infarction), and stroke. Stroke included overall stroke and main subtypes of stroke (e.g. ischemic stroke, including small vessel ischemic stroke, large vessel ischemic stroke and cardioembolic stroke as well as hemorrhagic stroke) (further details on data extraction can be found in Data S2).

Data synthesis and evaluation of robustness

Based on the extracted information, we presented the basic characteristics of the identified MR analyses. Main findings were categorized by risk factor and risk factor categories. The robustness of the evidence (*robust, probable, suggestive* and *insufficient* evidence) was assessed through a-priori defined criteria (Supplementary Figure S1 – Data S3) [6] based on previous recommendations [7]. We grouped MR studies into polygenic (*trans*) MR studies (studies which use variants from multiple regions of the genome associated with the risk factor

of interest) and monogenic (*cis*) MR studies (studies using biological knowledge and variants from a single gene region associated with the risk factor of interest). For example, a MR analysis for C-reactive protein (CRP) can be monogenic and therefore conducted using variants in the CRP gene only or polygenic and therefore conducted using all independent genome-wide significant variants associated with CRP [7].

For polygenic MR studies, we based the evaluation on the results of the main MR analysis and the sensitivity analyses (e.g., MR-Egger, WM, MRPRESSO). The sensitivity analyses are used to check potential violations of the assumptions of the MR methodology. Evidence for causality was therefore considered stronger when a sensitivity analysis was reported and was supportive of the main analysis findings as MR investigation that do not perform one or more sensitivity methods may be viewed as having incomplete evidence. More specifically, the associations were considered as having *robust* evidence for causality when all methods had concordant direction of effect estimates and both the main analysis and at least one sensitivity analysis achieved statistical significance (P<0.05). When studies adjusted their results for multiple testing, we used the p-value threshold after the adjustment to define statistical significance, otherwise we used nominal significance level (P<0.05). When a p-value was not reported for the main MR estimate, we calculated it using the effect size and standard error. Also, when studies also reported analyses excluding genetic variants with evidence of pleiotropy, we considered those as the main analysis as they better account for the MR assumptions. The term "robust" refers to evidence of causality for the studied associations, not the quality of the analysis. An association was supported by *probable* evidence for causality when at least one method (main or sensitivity analysis) achieved statistical significance and the direction of the effect estimate was concordant in all methods. *Suggestive* evidence for causality was achieved when at least one method had a statistically significant p-value, but the direction of the effect

estimates differed between methods. Associations that presented non-significant p-values for both the main analysis and sensitivity analyses were classified as *insufficient* evidence for causality. Polygenic MR studies which did not report any sensitivity analyses were *non-evaluable* based on the above grading scheme which focuses on evaluating the robustness for causality of the studied associations.

Monogenic MR studies included MR analyses examining single SNPs or single gene regions to define the risk factor (instrumental variable of interest).. Most of these studies could not preform sensitivity analyses as the number of genetic variants was small. We assessed the robustness of these results based on whether the authors perform also colocalization analysis[8]. Colocalization assesses whether the same genetic variant (or variants) influences two traits and is useful when MR is based on a single gene region [7].

Finally, we further assessed the reporting of all MR studies using the Strengthening the Reporting of Mendelian Randomization Studies (STROBE-MR) Guidelines [4,5].

All statistical analyses were done with R 4.1.0.

Results

Eligible studies

The literature search yielded 3,980 papers of which 586 were evaluated in full text and, of them, 391 publications were deemed eligible (see full list in Supplementary Files S1 and S2 respectively). The majority of studies were published from 2018 onwards (Supplementary Figure S2).

193 Description of study characteristics

194 Of 391 MR publications, 317 studied CAD as the outcome of interest, 175 stroke and 102 both
195 outcomes. Overall, the 391 publications included 2,725 different MR analyses examining 535
196 unique exposures, 482 in relation to CAD and 268 in relation to stroke, covering a broad range
197 of biomarkers, physical measurements, traits and diseases (Figure 2, Supplementary Table S1).
198 Many risk factors have been examined in multiple MR publications (Supplementary Table S2)
199 and the most commonly studied risk factors were LDL cholesterol for CAD (26 papers) and
200 body mass index (BMI) for stroke (10 papers).

201

202 There were 2,122 polygenic MR analyses and 596 monogenic MR analyses. The median
203 number of SNPs used to genetically predict the risk factor of interest in polygenic MR studies
204 was 13 ranging from two to 3,188 SNPs (Supplementary Table S3). The median sample size
205 for the exposure genetic analysis was 81,807 (with the smallest being 272 for phospholipase
206 A2 and the largest 1,887,658 for COVID-19 severity. GWAS summary statistics for the
207 exposures were derived from European (87.5%) and multi-ethnic (9.9%) ancestry populations.
208 Regarding the outcome, 65.9% of the associations were derived from European and 32.5%
209 from multi-ethnic populations. The vast majority of MR analyses used or included
210 CARDIoGRAMplusC4D (Coronary ARtery DIsease Genome wide Replication and Meta-
211 analysis (CARDIoGRAM) plus The Coronary Artery Disease (C4D) Genetics consortium)
212 when CAD was the outcome of interest (822 out of 1,478 MR analyses) and MEGASTROKE
213 when stroke was the outcome of interest (914 out of 1,214 MR analyses). Only 65 MR analyses
214 were based on one sample MR designs.

215

216 Evaluation of the robustness of causal associations

Supplementary Table S4 lists the main characteristics of each eligible MR analysis and its subsequent grading category based on the robustness of evidence, while supplementary Table S5 summarizes the grading categories. Out of the 2,122 polygenic MR associations, 20 analyses were based on two genetic variants (in different gene regions) and could not perform sensitivity analysis. From the remaining 2,102 MR analyses, 1,479 (70%) presented results on both the main and at least one sensitivity analysis and were eligible for evaluation. IVW was the main analysis in the majority of associations examined (N=1,931, 92%). Of the 1,530 associations reporting main and sensitivity analyses, we found 141 *robust* associations (median N_{SNPs}=110), 353 *probable* associations (median N_{SNPs}=71), 110 *suggestive* associations (median N_{SNPs}=73) and 926 associations with *insufficient* information (median N_{SNPs}=21). Overall, 276 MR analyses reported multivariable MR (MVMR) which examines multiple risk factors (exposures) simultaneously and estimates the independent causal effect of each of the risk factors. Genetically predicted BMI, smoking and lipid levels were common risk factors adjusted in MVMR.

There were also 596 monogenic MR associations examining genetic variants within a single gene region as instrumental variables for the risk factor of interest. Of them, 447 were based on a single genetic variant analysis only (Supplementary Figure S3). Among the 596 associations, 219 reported statistically significant results in the main analysis. Of them, 24 analyses performed colocalization analyses with the outcome of interest, of which two found evidence for colocalization between the risk factor and the outcome (Islet cell autoantigen 1-like protein (ICA1L) for stroke and neurokinin 3 receptor (NK3R) for CHD).

A graphical overview of the robustness of the evidence per exposure category and CVD group is presented in Figure 3. The exposure category with the most *robust* associations was anthropometry (N=28), followed by lipids and lipoproteins (N=24). There were 141 *robust* polygenic MR associations pertaining to 53 different risk factors as illustrated in detail in Supplementary Table S6 and Figure 4 (35 risk factors for CAD and 22 risk factors for stroke [9-76]. Almost all studies that showed robust evidence of association between an exposure and CAD or stroke had a good reporting score for MR-STROBE reporting, with the exception of three studies (Supplementary Table S7) [60,71,73]. Apart from conventional cardiovascular risk factors such as blood pressure traits, cholesterol levels, type 2 diabetes, obesity and smoking, *robust* positive associations were observed between genetically predicted calcium (OR_{IVW}: 1.66, 95%CI: 1.12-1.81), lymphocyte count (OR_{IVW}: 1.09, 95%CI:1.02-1.16), colony stimulating factor 1 (OR_{IVW}: 1.19, 95%CI: 1.08-1.30) and omega 6 fatty acid levels (OR_{IVW}: 1.21, 95%CI: 1.12-1.31) with CAD. Protective (inverse) associations were also observed for genetically predicted height (OR_{IVW}: 0.84, 95%CI:0.78–0.90), forced vital capacity (FVC) (OR_{IVW}: 0.64, 95%CI: 0.46 - 0.88), sex hormone binding globulin (OR_{IVW}: 0.79, 95%CI: 0.7-0.91) and interleukin 6 receptor (IL6R) (OR_{IVW}: 0.9, 95%CI:0.85-0.95). For stroke *robust* associations were observed for genetic predisposition to several thrombotic factors, apolipoprotein B (OR_{IVW}: 1.14, 95%CI: 1.07-1.22) and urinary sodium excretion (OR_{IVW}: 1.6, 95%CI: 1.12-2.3) (positive) and IL6R (OR_{IVW}: 0.92, 95%CI: 0.85 - 0.98) and transferrin (OR_{IVW}: 0.82, 95%CI: 0.70 - 0.96) (inverse).

Of the robust associations, 24 reported multivariable MR analyses adjusting for potential mediators. Of them, two associations were attenuated to the null: linoleic acid and stroke after adjusting for LDL cholesterol and FVC and CHD after adjusting for height. The rest of

multivariable analyses attenuated the estimates, suggesting different levels of mediation; however, statistical significance was retained (Supplementary Table S8).

Discussion

In this systematic review, we summarized the evidence for associations between genetic predisposition to 535 risk factors and CAD or stroke examined in 391 publications covering 2,725 MR associations. Using a set of predefined criteria, we found *robust* evidence for causality between 35 distinct risk factors and CAD and between 22 risk factors and stroke. For CAD, these included the well-established cardiovascular risk factors such as blood pressure, type 2 diabetes, obesity, smoking and cholesterol levels but also, anthropometry and physical measurements (height, birth weight, muscular strength and FVC) and several biomarkers (leucocyte count, serum calcium, IL6R signaling, protein C, omega 6 fatty acid levels, sex hormone binding globulin). Stroke showed a somewhat different profile of associations, with evidence for causal effect for thrombotic risk factors (factor VII, factor XI, platelet count, vitamin K), iron and inflammatory biomarkers (iron, transferrin, transferrin saturation, IL6R), and blood pressure.

This large body of published MR analyses highlighted several reporting limitations also observed in a previous systematic review of MR studies on cancer outcomes [77]. Approximately half of the associations included sensitivity analyses, which are important to assess the assumptions of the method and therefore the robustness of the results. The lack of sensitivity analyses was often because studies were published early, prior to the availability of MR sensitivity methods, or because they were monogenic (single-gene) or single variant MR studies where sensitivity analyses were not feasible due to the small number of IVs. In the latter

288 case of MR studies, colocalization can be used to investigate whether the exposure and the
289 outcome share a causal variant in the genetic region, but it was rarely performed in the
290 examined MR studies. Again, this may be due to the fact that colocalization was only suggested
291 recently as an additional method to support monogenic MR investigations and a large
292 proportion of these studies were published earlier. Genetic variants typically explain only a
293 small proportion of the variation in the relevant exposure of interest and as a result, low
294 statistical power is common in MR studies. MR studies examined here rarely reported power
295 estimates and/or the variance explained by IVs and therefore it was difficult to conclude
296 whether non-significant associations were true null findings [78]. The recent publication of MR
297 reporting guidelines (STROBE-MR) statement should improve the reporting standards of MR
298 studies and further enhance the robustness and interpretability of MR findings [5]. Finally, MR
299 investigations are dependent on the primary GWAS sources, the quality of which was not
300 assessed in this work. However, underlying GWAS quality are unlikely to lead to false positive
301 results.

302

303 A considerable proportion of the studies provided supporting evidence for causal associations
304 between the so-called conventional CVD risk factors and CVD events. We identified studies
305 with *robust* evidence for causal associations between genetically predicted LDL cholesterol,
306 HDL cholesterol, triglycerides, apolipoprotein B, blood pressure, type 2 diabetes, and glycemic
307 traits such as HbA1c and insulin resistance with CAD and stroke. This supports the extensive
308 evidence from traditional epidemiological studies, experimental studies, and RCTs examining
309 these risk factors. However, MR provided additional valuable information such as examination
310 of comparative effects between correlated risk factors, estimation of non-linear effects and
311 interactions with other factors. For example, multivariable MR analyses on several lipids and
312 lipoproteins highlighted the central role of apolipoprotein B compared to other lipids in

313 ischemic stroke [53]. Similarly, the MR paradigm generated evidence supporting an effect of
314 midlife blood pressure on later life CAD risk independent of later life blood pressure [41].

315

316 Measures of anthropometry have also been extensively studied in the MR context in relation
317 to CAD and stroke. Beyond BMI, which showed *robust* causal associations with CAD and
318 stroke, higher height was also highlighted as a potentially causal risk factor for CAD.
319 Genetically predicted height also mediated at least of the association between lung function
320 measured by FVC and CHD and stroke. Although several observational studies have reported
321 a protective role of short height for CVD, the magnitude of this association has been
322 controversial [79] and the mechanisms underlying this inverse association are not well understood
323 [80]. One proposed explanation is that shorter individuals have on average smaller vessel
324 diameter, which can lead to increased arterial occlusive events [81,82]. There was also both *robust*
325 and *probable* evidence for a protective association between higher birthweight and CAD and
326 stroke respectively supporting the fetal developmental origins of CVD [26,83]. Interestingly,
327 further investigation into the fetal and/or maternal components of instrumental effects on birth
328 weight showed *robust* evidence between lower birthweight, by maternal rather than fetal
329 genome, and stroke and its subtypes in later life [51].

330

331 Lifestyle is an important area of MR research in CVD as RCTs are often inappropriate or
332 unfeasible and evidence stemming from MR is vital to support causality. Smoking behavior
333 showed *robust* evidence for causal association with CVD in agreement with overwhelming
334 evidence from observational epidemiology [84]. Coffee, alcohol consumption and sleep duration
335 also showed *probable* associations [48,85-90]. Observational epidemiology has often suggested a
336 possible protective effect of moderate alcohol consumption on CVD. Dose response MR

analyses did not support this conclusion but they found evidence of a dose–response relationship between alcohol and risk of stroke [91]. Educational attainment was reported to have a protective role for both CAD and stroke, exhibiting *robust* evidence of association in MR studies [15,28,33]. Traditional observational studies and MR mediation analyses have shown that BMI, systolic blood pressure, and smoking behavior mediate a substantial proportion of the protective effect of education on the risk of CVD outcomes [92,93]. Despite the research interest on diet and CVD, there were few *robust* or *probable* associations between nutrients or dietary traits and CVD outcomes. This is in partly expected due to the weak genetic instruments on some nutrients and other dietary variables (few SNPs available to instrument dietary traits) and the small heritable components of many dietary traits both leading to under-powered MR studies.

Many MR analyses concentrate on the causal association between biomarker levels and CVD to identify novel treatment targets for the disease. Of them, genetically-predicted plasma soluble IL6 receptor (sIL6R), an IL6 signaling biomarker, showed *robust* evidence for an inverse association with CAD [36] and stroke [60] supporting a key role of inflammation in CVD which has also supportive RCT evidence [94]. Thrombotic factors are implicated in the coagulation cascade and along with inflammatory factors are contributing to the suppression of a pathogen entering in the host, a mechanism termed as immunothrombosis. The aberrant activation of immunothrombosis has been associated with increased risk for myocardial infarction, stroke and venous thromboembolism [95]. This association is supported by MR evidence. A *robust* positive association was observed between vitamin K and large vessel stroke as well as between two enzymes of the coagulation cascade (i.e., factor XI and factor VII) and ischemic stroke [19,22,30]. In contrast to thrombotic factors, Protein C, also known as factor XIX, a zymogen that inactivates thrombotic enzymes, showed evidence for an inverse

causal effect with CAD [14]. In concordance with meta-analyses of RCTs for calcium supplementation [96,97], MR evidence supported a causal association between higher serum calcium levels and increased CAD risk [17,98-100]. Circulating calcium levels are thought to increase CAD risk through vascular calcification [101,102] or via the upregulation of the coagulation pathway which in turn is associated with CVD risk [103]. Finally, iron, ferritin, and transferrin saturation, biomarkers of iron metabolism and intake [104], showed *robust* positive causal associations with risk of stroke in MR analyses and this effect was suggested to be driven by an increased risk of cardioembolic stroke [21]. The latter, along with the absence of an association with CAD, may indicate that the effect of iron on stroke is through thrombus formation rather than atherosclerosis [105].

Strengths and limitations

In the current systematic review, we summarized all previously published MR studies for all genetically determined exposures and their association with CVD risk. A clear categorization scheme and evaluation criteria were applied, to further examine the robustness credibility of the resulted associations. Other efforts to summarize the evidence of MR analyses on CVD risk have been performed in the past. However, they were either limited to specific exposures [106], or used a more narrative approach of presenting and assessing the MR results [107], while none performed a formal evaluation of the evidence. However, some limitations exist, which need to be acknowledged. Some relevant MR studies may have been missed through our search strategy, especially if the MR analysis was not the primary focus but only a supplementary analysis, which seems to be increasingly common in recent GWAS. In the absence of comprehensive MR guidelines, we based our evaluation of the evidence of causality adapting a set of previously proposed criteria. This approach did not allow us to investigate MR studies

386 presenting only main analysis without sensitivity analyses. Sensitivity analyses increase the
387 credibility of the findings as they test various MR assumptions. However, many studies did not
388 present those as they were published earlier before such analyses were introduced in the
389 literature or were based on monogenic associations with a small number of SNPs which did
390 not allow sensitivity analyses. For the latter associations we based our evaluation on
391 availability of colocalization analysis which was again introduced only recently. Therefore, the
392 evaluation criteria for this systematic review were designed mainly for the assessment of the
393 evidence that resulted from the MR analyses and not for the assessment of the quality of the
394 studies. Although many studies included instrumental variables from the largest available
395 GWAS for the exposure traits, the SNPs explained a small percentage of the variance and
396 therefore some studies were underpowered. Finally, information on statistical power of the
397 instrument was often not reported, and therefore the grading scheme used could not distinguish
398 between MR analyses with robust evidence of lack of association or MR analyses which did
399 not present an association due to lack of power.

400

401 **Conclusions**

402 MR studies have contributed a large body of evidence into the causal association between risk
403 factors and CVD. Although many studies concentrated on CVD risk factors known to be
404 causally associated to CVD through RCTs, MR provided confirmation of previous
405 associations, supported evidence for potentially novel causal risk factors as well as refuted
406 several associations suggested by observational studies. Despite the plethora of MR
407 investigations in CVD, the highlighted associations with *robust* evidence for causality were
408 modest. Those risk factors concentrated around conventional risk factors for CVD,
409 inflammation and thrombotic factors and indices of anthropometry and showed a large overlap
410 between risk factors for CAD and stroke as well as highlighted the different risk factor profiles

between stroke subtypes. As GWAS investigations of exposures become larger, novel exposures are measured in epidemiological settings, and novel MR methodologies are published, the contribution of MR in establishing causal associations and prioritizing RCT is expected to grow further.

Sources of Funding

The project is co-financed by the European Regional Development Fund of the European Union and Greek national funds through 1. the Operational Program Competitiveness, Entrepreneurship and Innovation (EPAnEK), NSRF 2014-2020 (Project code MIS: ΟΠΣ 5047228)", and 2. the Operational Programme Epirus 2014–2020 of the Prefecture of Epirus (Project code MIS: ΗΠ1ΑΒ-0028180). The authors acknowledge generous support by National Institutes of Health (R01HL133932 and R01HL111362), Medical Research Council and National Institute for Health Research [grant number MC_PC_12025], the MRC Centre for Environment and Health (MR/S019669/1) and UK Dementia Research Institute (MC_PC_17114), which is supported by the Medical Research Council, the Alzheimer's Society and Alzheimer's Research UK.

Disclosures

The authors have no disclosures to report.

Supplemental Material:

Supplementary Information file: Data S1-S3, Figures S1-S3

Supplementary files and tables: Files S1-S2, Tables S1-S8

References

1. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, Barengo NC, Beaton AZ, Benjamin EJ, Benziger CP, et al. Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. *Journal of the American College of Cardiology*. 2020;76:2982-3021.
2. Tzoulaki I, Elliott P, Kontis V, Ezzati M. Worldwide Exposures to Cardiovascular Risk Factors and Associated Health Effects: Current Knowledge and Data Gaps. *Circulation*. 2016;133:2314-2333.
3. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008;27:1133-1163.
4. Larsson SC, Burgess S. Causal role of high body mass index in multiple chronic diseases: a systematic review and meta-analysis of Mendelian randomization studies. *BMC Medicine*. 2021;19:320.
5. Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, VanderWeele TJ, Higgins JPT, Timpson NJ, Dimou N, et al. Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement. *Jama*. 2021;326:1614-1621.
6. Markozannes G, Kanellopoulou A, Dimopoulou O, Kosmidis D, Zhang X, Wang L, Theodoratou E, Gill D, Burgess S, Tsilidis KK. Systematic review of Mendelian randomization studies on risk of cancer. *BMC Med*. 2022;20:41.
7. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, Hartwig FP, Holmes MV, Minelli C, Relton CL, et al. Guidelines for performing Mendelian randomization investigations. *Wellcome open research*. 2019;4:186.
8. Kanduri C, Bock C, Gundersen S, Hovig E, Sandve GK. Colocalization analyses of genomic elements: approaches, recommendations and challenges. *Bioinformatics*. 2019;35:1615-1624.
9. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, Mead D, Bouman H, Riveros-Mckay F, Kostadima MA, et al. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell*. 2016;167:1415-1429.e1419.
10. Au Yeung SL, Lin SL, Li AM, Schooling CM. Birth weight and risk of ischemic heart disease: A Mendelian randomization study. *Sci Rep*. 2016;6:38420.
11. White J, Swerdlow DI, Preiss D, Fairhurst-Hunter Z, Keating BJ, Asselbergs FW, Sattar N, Humphries SE, Hingorani AD, Holmes MV. Association of Lipid Fractions With Risks for Coronary Artery Disease and Diabetes. *JAMA Cardiol*. 2016;1:692-699.
12. Dale CE, Fatemifar G, Palmer TM, White J, Prieto-Merino D, Zabaneh D, Engmann JEL, Shah T, Wong A, Warren HR, et al. Causal Associations of Adiposity and Body Fat Distribution With Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus: A Mendelian Randomization Analysis. *Circulation*. 2017;135:2373-2388.
13. Ference BA, Kastelein JJP, Ginsberg HN, Chapman MJ, Nicholls SJ, Ray KK, Packard CJ, Laufs U, Brook RD, Oliver-Williams C, et al. Association of Genetic Variants Related to CETP Inhibitors and Statins With Lipoprotein Levels and Cardiovascular Risk. *Jama*. 2017;318:947-956.
14. Schooling CM, Zhong Y. Plasma levels of the anti-coagulation protein C and the risk of ischaemic heart disease. A Mendelian randomisation study. *Thromb Haemost*. 2017;117:262-268.
15. Tillmann T, Vaucher J, Okbay A, Pikhart H, Peasey A, Kubinova R, Pajak A, Tamosiunas A, Malyutina S, Hartwig FP, et al. Education and coronary heart disease: mendelian randomisation study. *Bmj*. 2017;358:j3542.

16. Xu L, Borges MC, Hemani G, Lawlor DA. The role of glycaemic and lipid risk factors in mediating the effect of BMI on coronary heart disease: a two-step, two-sample Mendelian randomisation study. *Diabetologia*. 2017;60:2210-2220.
17. Xu L, Lin SL, Schooling CM. A Mendelian randomization study of the effect of calcium on coronary artery disease, myocardial infarction and their risk factors. *Sci Rep*. 2017;7:42691.
18. Geng T, Smith CE, Li C, Huang T. Childhood BMI and Adult Type 2 Diabetes, Coronary Artery Diseases, Chronic Kidney Disease, and Cardiometabolic Traits: A Mendelian Randomization Analysis. *Diabetes Care*. 2018;41:1089-1096.
19. Gill D, Georgakis MK, Laffan M, Sabater-Lleal M, Malik R, Tzoulaki I, Veltkamp R, Dehghan A. Genetically Determined FXI (Factor XI) Levels and Risk of Stroke. *Stroke*. 2018;49:2761-2763.
20. Gill D, Monori G, Georgakis MK, Tzoulaki I, Laffan M. Genetically Determined Platelet Count and Risk of Cardiovascular Disease. *Arterioscler Thromb Vasc Biol*. 2018;38:2862-2869.
21. Gill D, Monori G, Tzoulaki I, Dehghan A. Iron Status and Risk of Stroke. *Stroke*. 2018;49:2815-2821.
22. Larsson SC, Traylor M, Markus HS. Circulating Vitamin K_1 Levels in Relation to Ischemic Stroke and Its Subtypes: A Mendelian Randomization Study. *Nutrients*. 2018;10.
23. Lv WQ, Zhang X, Fan K, Xia X, Zhang Q, Liu HM, Jiang BY, Zhang WD, Deng HW. Genetically driven adiposity traits increase the risk of coronary artery disease independent of blood pressure, dyslipidaemia, glycaemic traits. *Eur J Hum Genet*. 2018;26:1547-1553.
24. Sjaarda J, Gerstein H, Chong M, Yusuf S, Meyre D, Anand SS, Hess S, Paré G. Blood CSF1 and CXCL12 as Causal Mediators of Coronary Artery Disease. *J Am Coll Cardiol*. 2018;72:300-310.
25. Tikkanen E, Gustafsson S, Amar D, Shcherbina A, Waggott D, Ashley EA, Ingelsson E. Biological Insights Into Muscular Strength: Genetic Findings in the UK Biobank. *Sci Rep*. 2018;8:6451.
26. Zanetti D, Tikkanen E, Gustafsson S, Priest JR, Burgess S, Ingelsson E. Birthweight, Type 2 Diabetes Mellitus, and Cardiovascular Disease: Addressing the Barker Hypothesis With Mendelian Randomization. *Circ Genom Precis Med*. 2018;11:e002054.
27. Allara E, Morani G, Carter P, Gkatzionis A, Zuber V, Foley CN, Rees JMB, Mason AM, Bell S, Gill D, et al. Genetic Determinants of Lipids and Cardiovascular Disease Outcomes: A Wide-Angled Mendelian Randomization Investigation. *Circ Genom Precis Med*. 2019;12:e002711.
28. Carter AR, Gill D, Davies NM, Taylor AE, Tillmann T, Vaucher J, Wootton RE, Munafò MR, Hemani G, Malik R, et al. Understanding the consequences of education inequality on cardiovascular disease: mendelian randomisation study. *Bmj*. 2019;365:l1855.
29. Censin JC, Peters SAE, Bovijn J, Ferreira T, Pulit SL, Mägi R, Mahajan A, Holmes MV, Lindgren CM. Causal relationships between obesity and the leading causes of death in women and men. *PLoS Genet*. 2019;15:e1008405.
30. de Vries PS, Sabater-Lleal M, Huffman JE, Marten J, Song C, Pankratz N, Bartz TM, de Haan HG, Delgado GE, Eicher JD, et al. A genome-wide association study identifies new loci for factor VII and implicates factor VII in ischemic stroke etiology. *Blood*. 2019;133:967-977.
31. Fang X, Zuo J, Zhou J, Cai J, Chen C, Xiang E, Li H, Cheng X, Chen P. Childhood obesity leads to adult type 2 diabetes and coronary artery diseases: A 2-sample mendelian randomization study. *Medicine (Baltimore)*. 2019;98:e16825.
32. Gan W, Bragg F, Walters RG, Millwood IY, Lin K, Chen Y, Guo Y, Vaucher J, Bian Z, Bennett D, et al. Genetic Predisposition to Type 2 Diabetes and Risk of Subclinical Atherosclerosis and Cardiovascular Diseases Among 160,000 Chinese Adults. *Diabetes*. 2019;68:2155-2164.
33. Gill D, Efstathiadou A, Cawood K, Tzoulaki I, Dehghan A. Education protects against coronary heart disease and stroke independently of cognitive function: evidence from Mendelian randomization. *Int J Epidemiol*. 2019;48:1468-1477.
34. Leong A, Chen J, Wheeler E, Hivert MF, Liu CT, Merino J, Dupuis J, Tai ES, Rotter JI, Florez JC, et al. Mendelian Randomization Analysis of Hemoglobin A(1c) as a Risk Factor for Coronary Artery Disease. *Diabetes Care*. 2019;42:1202-1208.

35. Marouli E, Del Greco MF, Astley CM, Yang J, Ahmad S, Berndt SI, Caulfield MJ, Evangelou E, McKnight B, Medina-Gomez C, et al. Mendelian randomisation analyses find pulmonary factors mediate the effect of height on coronary artery disease. *Commun Biol*. 2019;2:119.
36. Rosa M, Chignon A, Li Z, Boulanger MC, Arsenault BJ, Bossé Y, Thériault S, Mathieu P. A Mendelian randomization study of IL6 signaling in cardiovascular diseases, immune-related disorders and longevity. *NPJ Genom Med*. 2019;4:23.
37. Valdes-Marquez E, Parish S, Clarke R, Stari T, Worrall BB, Hopewell JC. Relative effects of LDL-C on ischemic stroke and coronary disease: A Mendelian randomization study. *Neurology*. 2019;92:e1176-e1187.
38. Au Yeung SL, Schooling CM. Impact of urinary sodium on cardiovascular disease and risk factors: A 2 sample Mendelian randomization study. *Clin Nutr*. 2020.
39. Chen W, Wang S, Lv W, Pan Y. Causal associations of insulin resistance with coronary artery disease and ischemic stroke: a Mendelian randomization analysis. *BMJ Open Diabetes Res Care*. 2020;8.
40. Georgakis MK, Gill D, Webb AJS, Evangelou E, Elliott P, Sudlow CLM, Dehghan A, Malik R, Tzoulaki I, Dichgans M. Genetically determined blood pressure, antihypertensive drug classes, and risk of stroke subtypes. *Neurology*. 2020;95:e353-e361.
41. Gill D, Georgakis MK, Zuber V, Karhunen V, Burgess S, Malik R, Dichgans M. Genetically Predicted Midlife Blood Pressure and Coronary Artery Disease Risk: Mendelian Randomization Analysis. *J Am Heart Assoc*. 2020;9:e016773.
42. Hu X, Zhuang XD, Mei WY, Liu G, Du ZM, Liao XX, Li Y. Exploring the causal pathway from body mass index to coronary heart disease: a network Mendelian randomization study. *Ther Adv Chronic Dis*. 2020;11:2040622320909040.
43. Lee SH, Lee JY, Kim GH, Jung KJ, Lee S, Kim HC, Jee SH. Two-Sample Mendelian Randomization Study of Lipid levels and Ischemic Heart Disease. *Korean Circ J*. 2020;50:940-948.
44. Liao LZ, Li WD, Liu Y, Li JP, Zhuang XD, Liao XX. Exploring the causal pathway from omega-6 levels to coronary heart disease: A network Mendelian randomization study. *Nutr Metab Cardiovasc Dis*. 2020;30:233-240.
45. Liu HM, Zhang Q, Shen WD, Li BY, Lv WQ, Xiao HM, Deng HW. Sarcopenia-related traits and coronary artery disease: a bi-directional Mendelian randomization study. *Aging (Albany NY)*. 2020;12:3340-3353.
46. Marini S, Merino J, Montgomery BE, Malik R, Sudlow CL, Dichgans M, Florez JC, Rosand J, Gill D, Anderson CD. Mendelian Randomization Study of Obesity and Cerebrovascular Disease. *Ann Neurol*. 2020;87:516-524.
47. Si S, Tewara MA, Li Y, Li W, Chen X, Yuan T, Liu C, Li J, Wang B, Li H, et al. Causal Pathways from Body Components and Regional Fat to Extensive Metabolic Phenotypes: A Mendelian Randomization Study. *Obesity (Silver Spring)*. 2020;28:1536-1549.
48. Thom CS, Ding Z, Levin MG, Damrauer SM, Lee KM, Lynch J, Chang KM, Tsao PS, Cho K, Wilson PWF, et al. Genetic determinants of increased body mass index mediate the effect of smoking on increased risk for type 2 diabetes but not coronary artery disease. *Hum Mol Genet*. 2020;29:3327-3337.
49. Thomas DG, Wei Y, Tall AR. Lipid and Metabolic Syndrome Traits in Coronary Artery Disease: A Mendelian Randomization Study. *J Lipid Res*. 2020.
50. Verkouter I, de Mutsert R, Smit RAJ, Trompet S, Rosendaal FR, van Heemst D, Willems van Dijk K, Noordam R. The contribution of tissue-grouped BMI-associated gene sets to cardiometabolic-disease risk: a Mendelian randomization study. *Int J Epidemiol*. 2020;49:1246-1256.
51. Wang T, Tang Z, Yu X, Gao Y, Guan F, Li C, Huang S, Zheng J, Zeng P. Birth Weight and Stroke in Adult Life: Genetic Correlation and Causal Inference With Genome-Wide Association Data Sets. *Front Neurosci*. 2020;14:479.

52. Ye D, Huang H, Wu DJH, Zhang W, Zhou F, Qian Y, Zheng J, Mao Y. Association Between Circulating Linoleic Acid and Risk of Ischemic Stroke. *Front Genet*. 2020;11:582623.
53. Yuan S, Tang B, Zheng J, Larsson SC. Circulating Lipoprotein Lipids, Apolipoproteins and Ischemic Stroke. *Ann Neurol*. 2020;88:1229-1236.
54. Benn M, Emanuelsson F, Tybjærg-Hansen A, Nordestgaard BG. Impact of high glucose levels and glucose lowering on risk of ischaemic stroke: a Mendelian randomisation study and meta-analysis. *Diabetologia*. 2021;64:1492-1503.
55. Chan, II, Kwok MK, Schooling CM. The total and direct effects of systolic and diastolic blood pressure on cardiovascular disease and longevity using Mendelian randomisation. *Sci Rep*. 2021;11:21799.
56. Cupido AJ, Asselbergs FW, Natarajan P, Ridker PM, Hovingh GK, Schmidt AF. Dissecting the IL-6 pathway in cardiometabolic disease: A Mendelian randomization study on both IL6 and IL6R. *British journal of clinical pharmacology*. 2021.
57. Fatumo S, Karhunen V, Chikowore T, Sounkou T, Udosen B, Ezenwa C, Nakabuye M, Soremekun O, Daghlas I, Ryan DK, et al. Metabolic Traits and Stroke Risk in Individuals of African Ancestry: Mendelian Randomization Analysis. *Stroke*. 2021;52:2680-2684.
58. Georgakis MK, Harshfield EL, Malik R, Franceschini N, Langenberg C, Wareham NJ, Markus HS, Dichgans M. Diabetes Mellitus, Glycemic Traits, and Cerebrovascular Disease: A Mendelian Randomization Study. *Neurology*. 2021;96:e1732-e1742.
59. Gill D, Zuber V, Dawson J, Pearson-Stuttard J, Carter AR, Sanderson E, Karhunen V, Levin MG, Wootton RE, Klarin D, et al. Risk factors mediating the effect of body mass index and waist-to-hip ratio on cardiovascular outcomes: Mendelian randomization analysis. *Int J Obes (Lond)*. 2021;45:1428-1438.
60. Higgins H, Mason AM, Larsson SC, Gill D, Langenberg C, Burgess S. Estimating the Population Benefits of Blood Pressure Lowering: A Wide-Angled Mendelian Randomization Study in UK Biobank. *J Am Heart Assoc*. 2021;10:e021098.
61. Levin MG, Klarin D, Assimes TL, Freiberg MS, Ingelsson E, Lynch J, Natarajan P, O'Donnell C, Rader DJ, Tsao PS, et al. Genetics of Smoking and Risk of Atherosclerotic Cardiovascular Diseases: A Mendelian Randomization Study. *JAMA Netw Open*. 2021;4:e2034461.
62. Li MY, Kwok MK, Schooling CM. Investigating Effects of Plasma Apolipoprotein E on Ischemic Heart Disease Using Mendelian Randomization Study. *Nutrients*. 2021;13.
63. Lu Y, Wang Z, Georgakis MK, Lin H, Zheng L. Genetic Liability to Depression and Risk of Coronary Artery Disease, Myocardial Infarction, and Other Cardiovascular Outcomes. *J Am Heart Assoc*. 2021;10:e017986.
64. Park S, Lee S, Kim Y, Lee Y, Kang M, Kim K, Kim Y, Han S, Lee H, Lee J, et al. Causal Effects of Serum Levels of n-3 or n-6 Polyunsaturated Fatty Acids on Coronary Artery Disease: Mendelian Randomization Study. *Nutrients*. 2021;13.
65. Tian D, Zhang L, Zhuang Z, Huang T, Fan D. A Mendelian randomization analysis of the relationship between cardioembolic risk factors and ischemic stroke. *Sci Rep*. 2021;11:14583.
66. Yuan S, Baron JA, Michaëlsson K, Larsson SC. Serum calcium and 25-hydroxyvitamin D in relation to longevity, cardiovascular disease and cancer: a Mendelian randomization study. *NPJ Genom Med*. 2021;6:86.
67. Zhao JV, Schooling CM. Genetically predicted sex hormone binding globulin and ischemic heart disease in men and women: a univariable and multivariable Mendelian randomization study. *Sci Rep*. 2021;11:23172.
68. Au Yeung SL, Borges MC, Lawlor DA, Schooling CM. Impact of lung function on cardiovascular diseases and cardiovascular risk factors: a two sample bidirectional Mendelian randomisation study. *Thorax*. 2022;77:164-171.
69. Chen K, Zhuang Z, Shao C, Zheng J, Zhou Q, Dong E, Huang T, Tang YD. Roles of Cardiometabolic Factors in Mediating the Causal Effect of Type 2 Diabetes on Cardiovascular

Diseases: A Two-Step, Two-Sample Multivariable Mendelian Randomization Study. *Front Cardiovasc Med*. 2022;9:813208.
70. Frerich S, Malik R, Georgakis MK, Sinner MF, Kittner SJ, Mitchell BD, Dichgans M. Cardiac Risk Factors for Stroke: A Comprehensive Mendelian Randomization Study. *Stroke*. 2022;53:e130-e135.
71. Grace C, Hopewell JC, Watkins H, Farrall M, Goel A. Robust estimates of heritable coronary disease risk in individuals with type 2 diabetes. *Genet Epidemiol*. 2022;46:51-62.
72. Huang Y, Liu Y, Ma Y, Tu T, Liu N, Bai F, Xiao Y, Liu C, Hu Z, Lin Q, et al. Associations of Visceral Adipose Tissue, Circulating Protein Biomarkers, and Risk of Cardiovascular Diseases: A Mendelian Randomization Analysis. *Front Cell Dev Biol*. 2022;10:840866.
73. Linden AB, Clarke R, Hammami I, Hopewell JC, Guo Y, Whiteley WN, Lin K, Turnbull I, Chen Y, Yu C, et al. Genetic associations of adult height with risk of cardioembolic and other subtypes of ischemic stroke: A mendelian randomization study in multiple ancestries. *PLoS Med*. 2022;19:e1003967.
74. Si S, Hou L, Chen X, Li W, Liu X, Liu C, Li Y, Yuan T, Li J, Wang B, et al. Exploring the Causal Roles of Circulating Remnant Lipid Profile on Cardiovascular and Cerebrovascular Diseases: Mendelian Randomization Study. *Journal of epidemiology*. 2022;32:205-214.
75. Yang F, Hu T, Chen S, Wang K, Qu Z, Cui H. Low Intelligence Predicts Higher Risks of Coronary Artery Disease and Myocardial Infarction: Evidence From Mendelian Randomization Study. *Front Genet*. 2022;13:756901.
76. Yang F, Lu Y, Chen S, Wang K, Hu T, Cui H. Sex-specific effect of serum urate levels on coronary heart disease and myocardial infarction prevention: A Mendelian randomization study. *Nutr Metab Cardiovasc Dis*. 2022;32:1266-1274.
77. Markozannes G, Kanellopoulou A, Dimopoulou O, Kosmidis D, Zhang X, Wang L, Theodoratou E, Gill D, Burgess S, Tsilidis KK. Systematic review of Mendelian randomization studies on risk of cancer. *BMC Medicine*. 2022;20:41.
78. Pierce BL, Ahsan H, Vanderweele TJ. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *International journal of epidemiology*. 2011;40:740-752.
79. Paajanen TA, Oksala NK, Kuukasjärvi P, Karhunen PJ. Short stature is associated with coronary heart disease: a systematic review of the literature and a meta-analysis. *Eur Heart J*. 2010;31:1802-1809.
80. Nüesch E, Dale C, Palmer TM, White J, Keating BJ, van Iperen EP, Goel A, Padmanabhan S, Asselbergs FW, Investigators E-N, et al. Adult height, coronary heart disease and stroke: a multi-locus Mendelian randomization meta-analysis. *International Journal of Epidemiology*. 2015;45:1927-1937.
81. Kortelainen ML, Särkioja T. Coronary atherosclerosis associated with body structure and obesity in 599 women aged between 15 and 50 years. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity*. 1999;23:838-844.
82. Lemos PA, Ribeiro EE, Perin MA, Kajita LJ, de Magalhães MA, Falcão JL, Filho AE, da Gama MN, Horta PE, Marchiori GG, et al. Angiographic segment size in patients referred for coronary intervention is influenced by constitutional, anatomical, and clinical features. *The international journal of cardiovascular imaging*. 2007;23:1-7.
83. Zeng P, Zhou X. Causal Association Between Birth Weight and Adult Diseases: Evidence From a Mendelian Randomization Analysis. *Front Genet*. 2019;10:618.
84. Mons U, Müezzinler A, Gellert C, Schöttker B, Abnet CC, Bobak M, de Groot L, Freedman ND, Jansen E, Kee F, et al. Impact of smoking and smoking cessation on cardiovascular events and mortality among older adults: meta-analysis of individual participant data from prospective cohort studies of the CHANCES consortium. *BMJ : British Medical Journal*. 2015;350:h1551.

85. Qian Y, Ye D, Wu DJ, Feng C, Zeng Z, Ye L, Zhu R, Zhang Z, Mao Y. Role of cigarette smoking in the development of ischemic stroke and its subtypes: a Mendelian randomization study. *Clin Epidemiol*. 2019;11:725-731.
86. Cai H, Liang J, Liu Z, Fang L, Zheng J, Xu J, Chen L, Sun W, Zhang H. Causal Effects of Sleep Traits on Ischemic Stroke and Its Subtypes: A Mendelian Randomization Study. *Nat Sci Sleep*. 2020;12:783-790.
87. Larsson SC, Burgess S, Mason AM, Michaëlsson K. Alcohol Consumption and Cardiovascular Disease: A Mendelian Randomization Study. *Circ Genom Precis Med*. 2020;13:e002814.
88. Larsson SC, Mason AM, Bäck M, Klarin D, Damrauer SM, Michaëlsson K, Burgess S. Genetic predisposition to smoking in relation to 14 cardiovascular diseases. *Eur Heart J*. 2020;41:3304-3310.
89. Lu H, Wu PF, Li RZ, Zhang W, Huang GX. Sleep Duration and Stroke: A Mendelian Randomization Study. *Frontiers in neurology*. 2020;11:976.
90. Qian Y, Ye D, Huang H, Wu DJH, Zhuang Y, Jiang X, Mao Y. Coffee Consumption and Risk of Stroke: A Mendelian Randomization Study. *Ann Neurol*. 2020;87:525-532.
91. Millwood IY, Walters RG, Mei XW, Guo Y, Yang L, Bian Z, Bennett DA, Chen Y, Dong C, Hu R, et al. Conventional and genetic evidence on alcohol and vascular disease aetiology: a prospective study of 500 000 men and women in China. *Lancet*. 2019;393:1831-1842.
92. 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;303:1159-1166.
93. Manrique-Garcia E, Sidorchuk A, Hallqvist J, Moradi T. Socioeconomic position and incidence of acute myocardial infarction: a meta-analysis. *Journal of epidemiology and community health*. 2011;65:301-309.
94. Ridker PM, Everett BM, Thuren T, MacFadyen JG, Chang WH, Ballantyne C, Fonseca F, Nicolau J, Koenig W, Anker SD, et al. Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. *N Engl J Med*. 2017;377:1119-1131.
95. Stark K, Massberg S. Interplay between inflammation and thrombosis in cardiovascular pathology. *Nature Reviews Cardiology*. 2021;18:666-682.
96. Mao PJ, Zhang C, Tang L, Xian YQ, Li YS, Wang WD, Zhu XH, Qiu HL, He J, Zhou YH. Effect of calcium or vitamin D supplementation on vascular outcomes: a meta-analysis of randomized controlled trials. *Int J Cardiol*. 2013;169:106-111.
97. Wang L, Manson JE, Sesso HD. Calcium intake and risk of cardiovascular disease: a review of prospective studies and randomized clinical trials. *American journal of cardiovascular drugs : drugs, devices, and other interventions*. 2012;12:105-116.
98. Larsson SC, Burgess S, Michaëlsson K. Association of Genetic Variants Related to Serum Calcium Levels With Coronary Artery Disease and Myocardial Infarction. *Jama*. 2017;318:371-380.
99. Zanetti D, Gustafsson S, Assimes TL, Ingelsson E. Comprehensive Investigation of Circulating Biomarkers and Their Causal Role in Atherosclerosis-Related Risk Factors and Clinical Events. *Circ Genom Precis Med*. 2020;13:e002996.
100. Zhou A, Morris HA, Hyppönen E. Health effects associated with serum calcium concentrations: evidence from MR-PheWAS analysis in UK Biobank. *Osteoporos Int*. 2019;30:2343-2348.
101. Kwak SM, Kim JS, Choi Y, Chang Y, Kwon MJ, Jung JG, Jeong C, Ahn J, Kim HS, Shin H, et al. Dietary intake of calcium and phosphorus and serum concentration in relation to the risk of coronary artery calcification in asymptomatic adults. *Arterioscler Thromb Vasc Biol*. 2014;34:1763-1769.
102. Shin S, Kim KJ, Chang HJ, Cho I, Kim YJ, Choi BW, Rhee Y, Lim SK, Yang WI, Shim CY, et al. Impact of serum calcium and phosphate on coronary atherosclerosis detected by cardiac computed tomography. *Eur Heart J*. 2012;33:2873-2881.

103. Reid IR, Gamble GD, Bolland MJ. Circulating calcium concentrations, vascular disease and mortality: a systematic review. *J Intern Med*. 2016;279:524-540.
104. Wish JB. Assessing iron status: beyond serum ferritin and transferrin saturation. *Clinical journal of the American Society of Nephrology : CJASN*. 2006;1 Suppl 1:S4-8.
105. Tang X, Fang M, Cheng R, Zhang Z, Wang Y, Shen C, Han Y, Lu Q, Du Y, Liu Y, et al. Iron-deficiency and estrogen are associated with ischemic stroke by up-regulating transferrin to induce hypercoagulability. *bioRxiv*. 2019:646109.
106. Riaz H, Khan MS, Siddiqi TJ, Usman MS, Shah N, Goyal A, Khan SS, Mookadam F, Krasuski RA, Ahmed H. Association Between Obesity and Cardiovascular Outcomes: A Systematic Review and Meta-analysis of Mendelian Randomization Studies. *JAMA Netw Open*. 2018;1:e183788.
107. Acosta JN, Szejko N, Falcone GJ. Mendelian Randomization in Stroke: A Powerful Approach to Causal Inference and Drug Target Validation. *Frontiers in Genetics*. 2021;12.

Figure legends

Figure 1: Flow chart of systematic literature search

Figure 2: Number of Mendelian randomization (MR) associations extracted from eligible publications according to different exposure categories for coronary artery disease (CAD) and stroke.

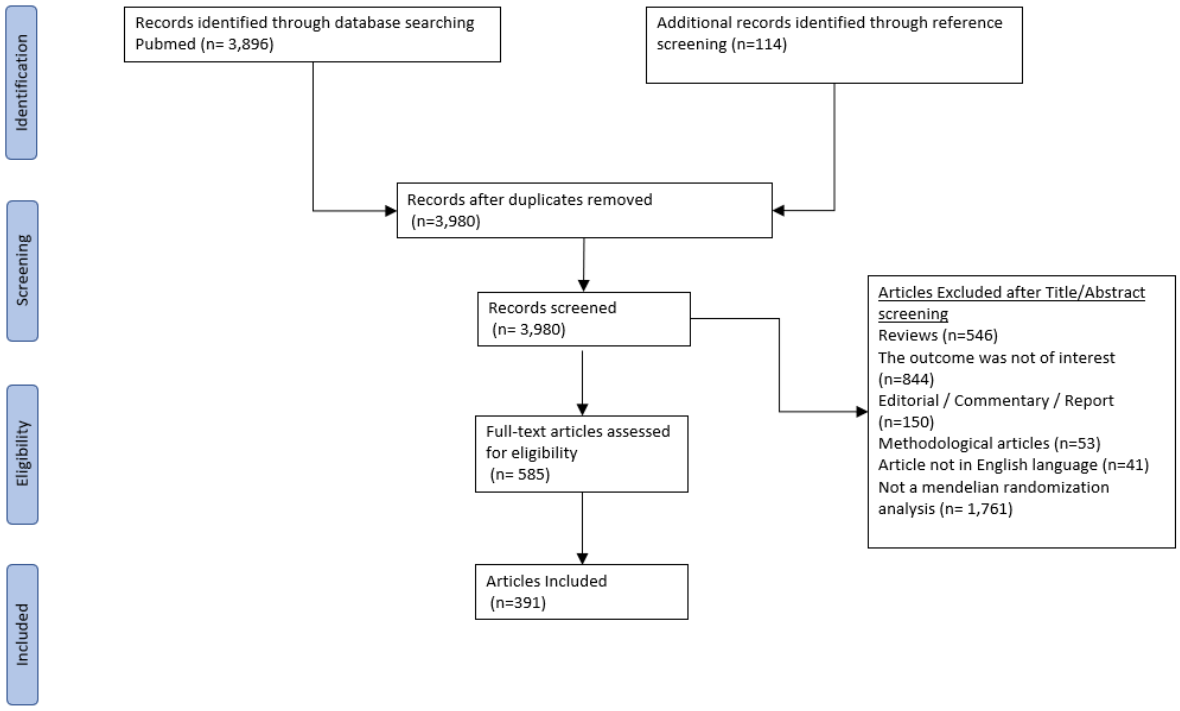
Figure 3: **Evidence map for eligible Mendelian randomization (MR) studies per exposure category for coronary artery disease (CAD) and stroke.** DNA methylation was not included in the diagram because of the limited number of analyses for the specific exposure category. Non-evaluable evidence level includes associations for which a sensitivity analysis was not feasible (e.g., single genetic variant analyses).

Figure 4: **Forest plot showing the identified *robust* associations between exposures and coronary artery disease (CAD) (A) and stroke (B).** When more than one study exhibited a *robust* association with coronary artery disease (CAD) or stroke, the most recent study with the largest sample size (exposure) was selected. The number in square brackets corresponds to the reference which examined the relevant exposure. Abbreviations: FVC: forced vital capacity, MI: Myocardial infarction, IS: Ischemic stroke; LVIS: Large vessel ischemic stroke, CIS: Cardioembolic ischemic stroke, Ref: Reference number

771 **Figure 1**

772

773



774

775

Figure 2

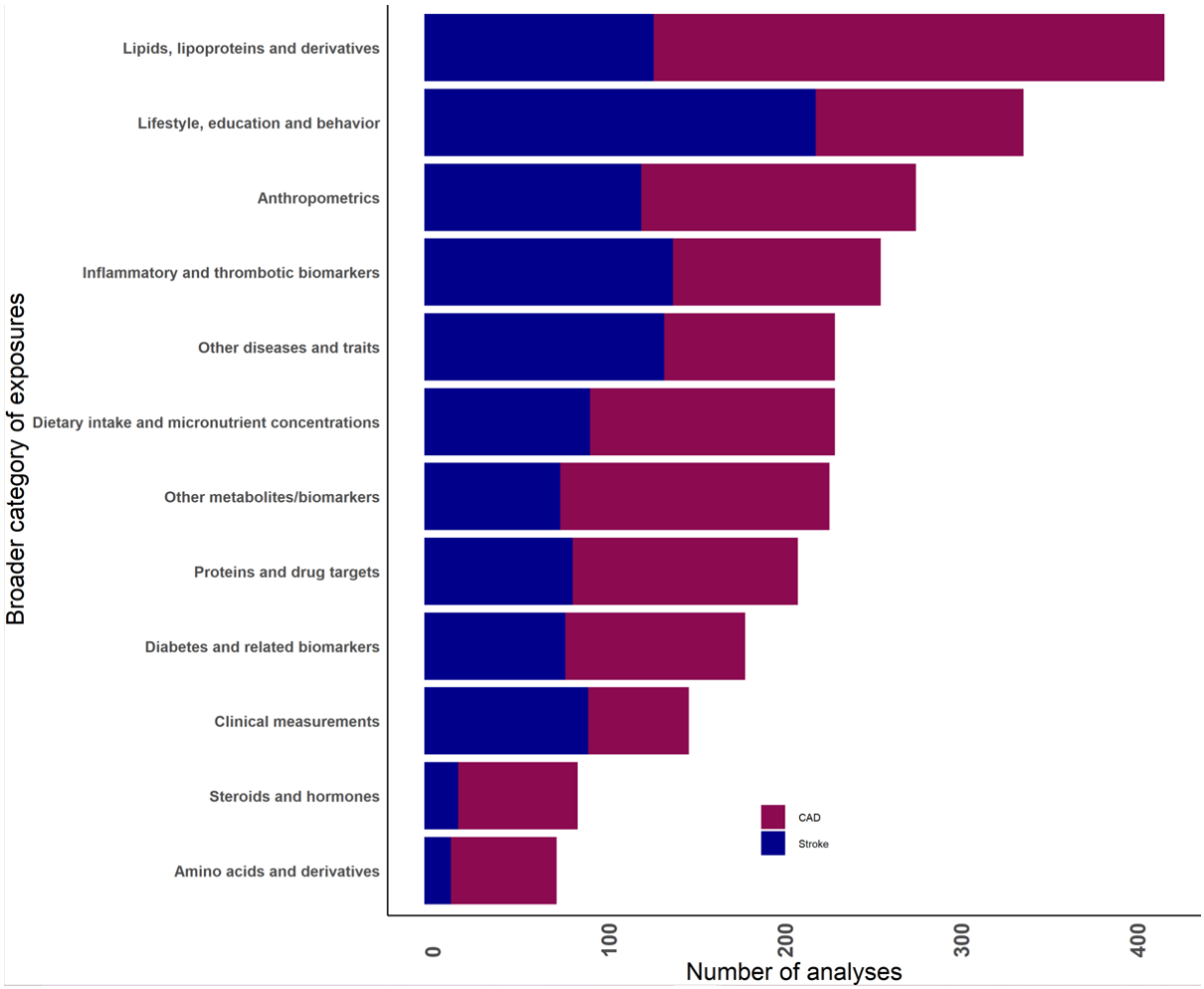

Lipids, lipoproteins and derivatives
Lifestyle, education and behavior
Anthropometrics
Inflammatory and thrombotic biomarkers
Other diseases and traits
Dietary intake and micronutrient concentrations
Other metabolites/biomarkers
Proteins and drug targets
Diabetes and related biomarkers
Clinical measurements
Steroids and hormones
Amino acids and derivatives
Broader category of exposures
CAD
Stroke
0
100
200
300
400
Number of analyses

Figure 3

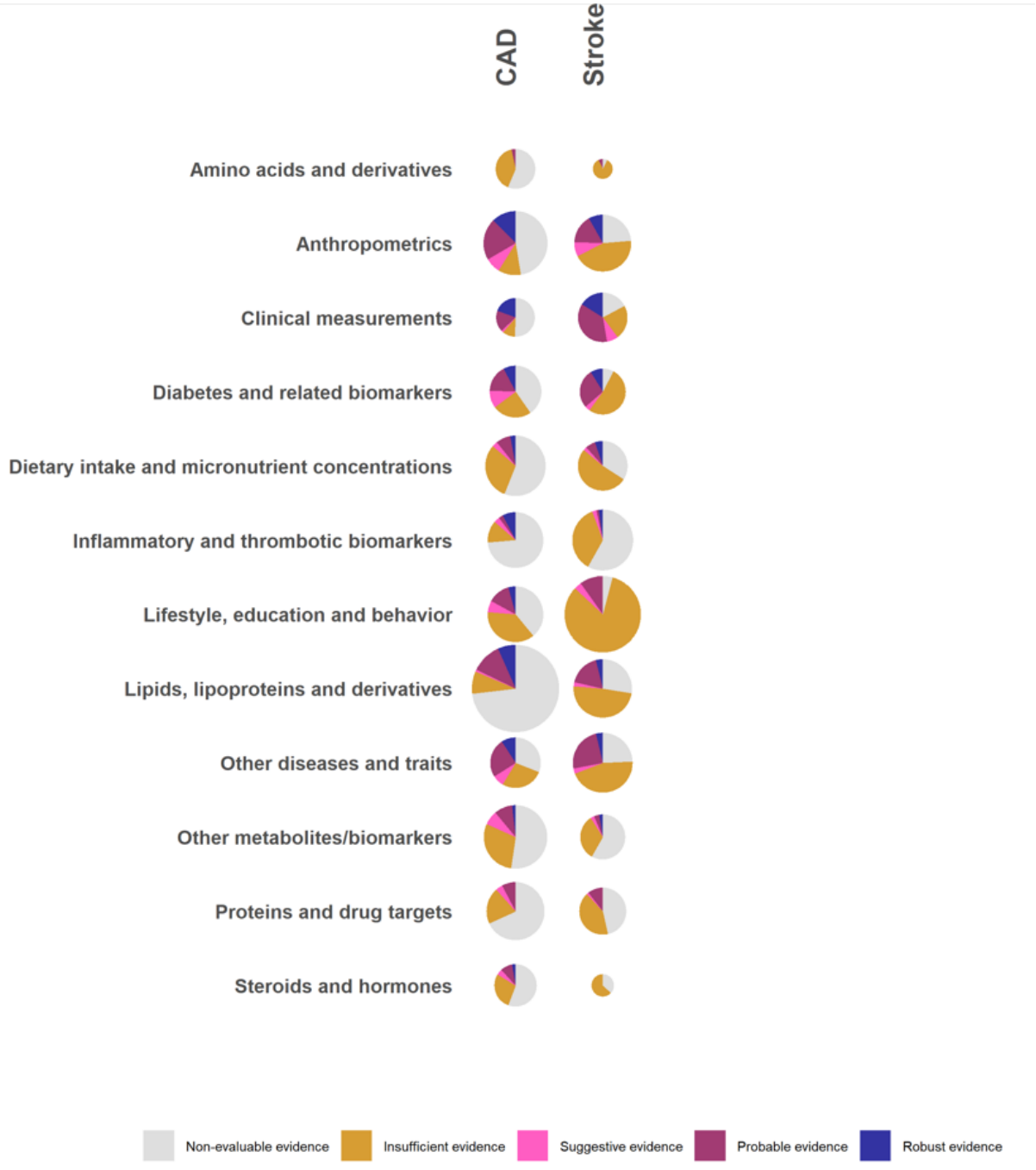

CAD
Stroke
Amino acids and derivatives
Anthropometrics
Clinical measurements
Diabetes and related biomarkers
Dietary intake and micronutrient concentrations
Inflammatory and thrombotic biomarkers
Lifestyle, education and behavior
Lipids, lipoproteins and derivatives
Other diseases and traits
Other metabolites/biomarkers
Proteins and drug targets
Steroids and hormones
Non-evaluable evidence
Insufficient evidence
Suggestive evidence
Probable evidence
Robust evidence

Figure 4 A

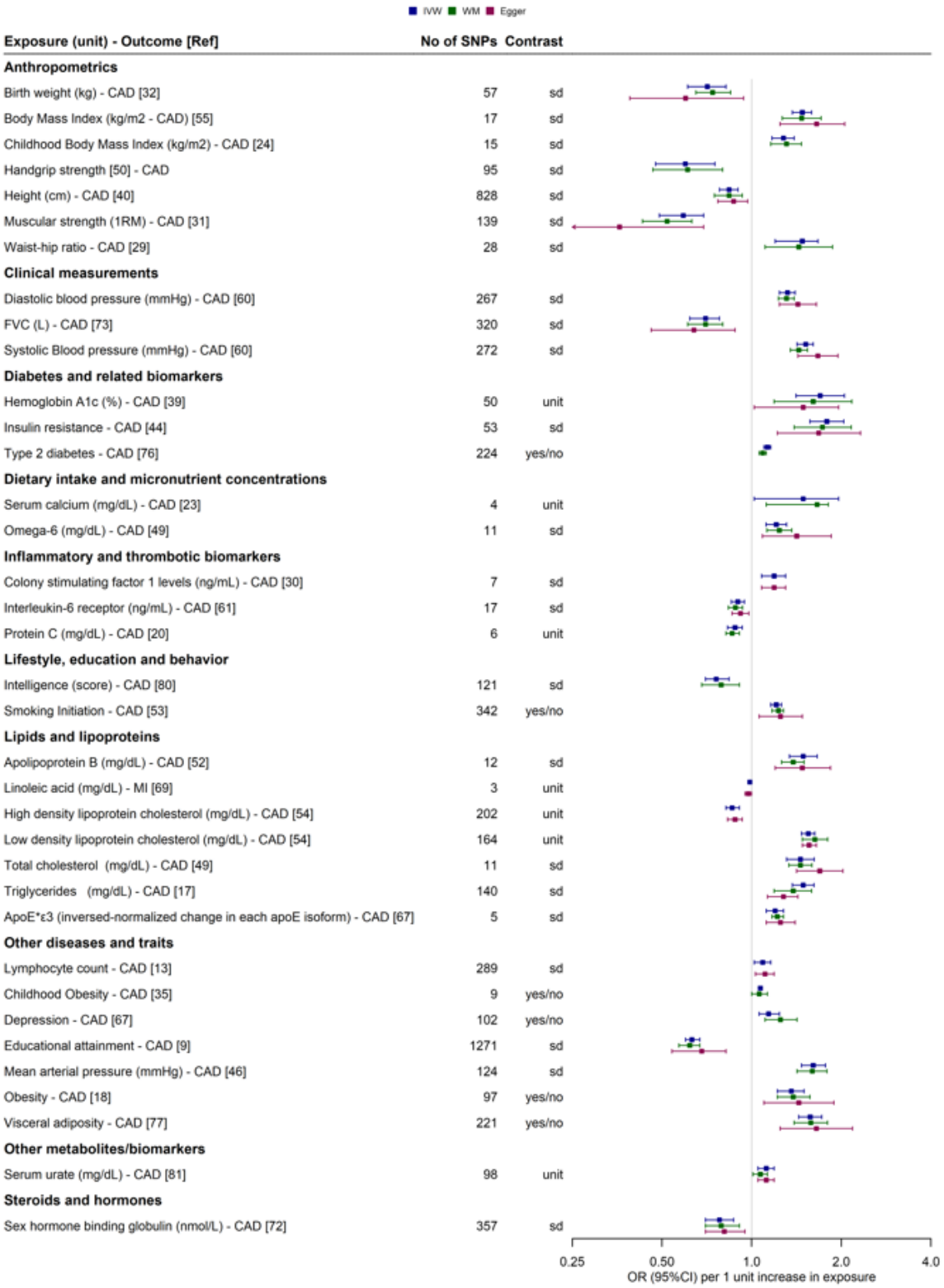
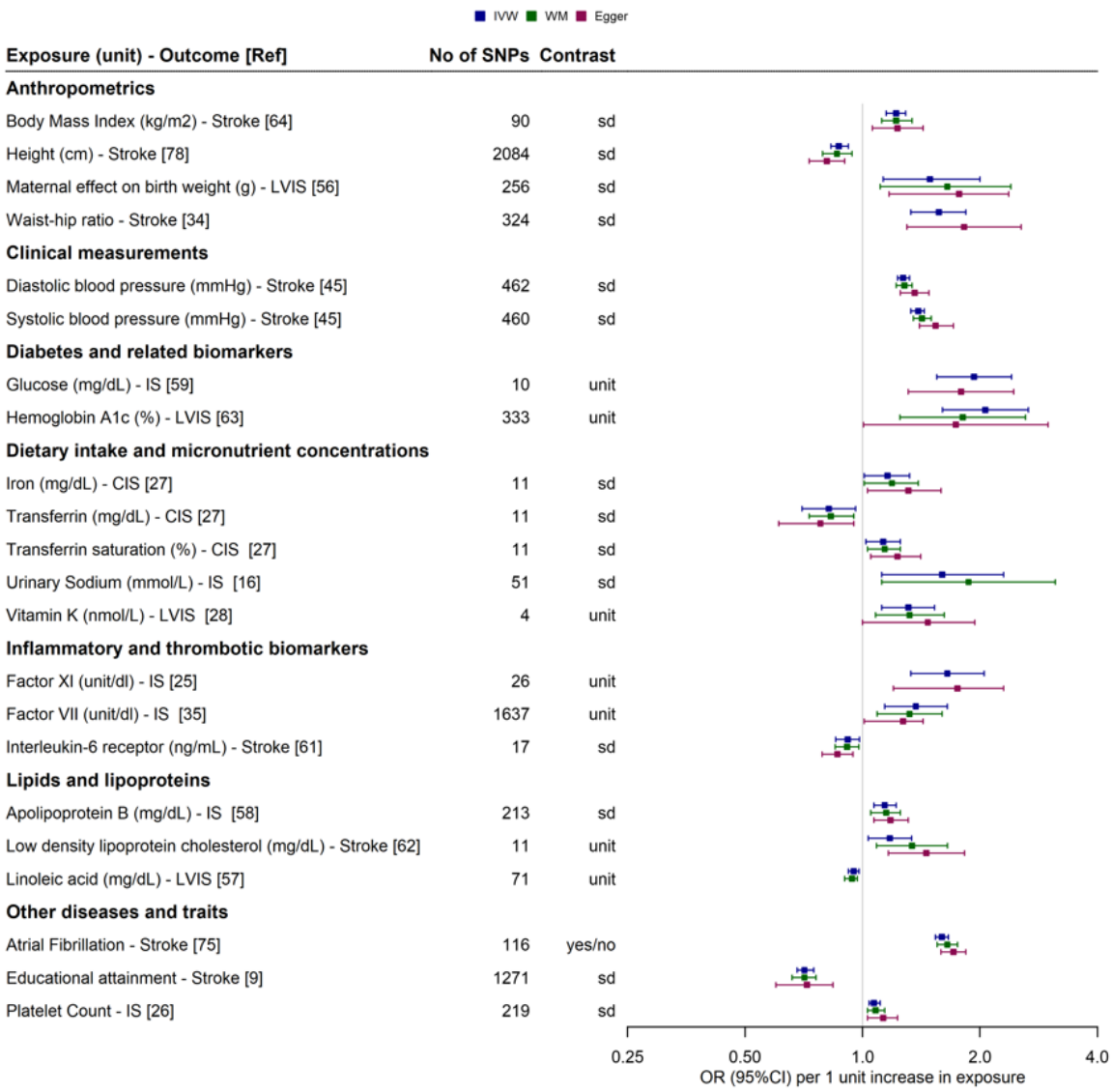

IVW WM Egger
Exposure (unit) - Outcome [Ref]
No of SNPs
Contrast
Anthropometrics
Birth weight (kg) - CAD [32] 57 sd
Body Mass Index (kg/m2 - CAD) [55] 17 sd
Childhood Body Mass Index (kg/m2) - CAD [24] 15 sd
Handgrip strength [50] - CAD 95 sd
Height (cm) - CAD [40] 828 sd
Muscular strength (1RM) - CAD [31] 139 sd
Waist-hip ratio - CAD [29] 28 sd
Clinical measurements
Diastolic blood pressure (mmHg) - CAD [60] 267 sd
FVC (L) - CAD [73] 320 sd
Systolic Blood pressure (mmHg) - CAD [60] 272 sd
Diabetes and related biomarkers
Hemoglobin A1c (%) - CAD [39] 50 unit
Insulin resistance - CAD [44] 53 sd
Type 2 diabetes - CAD [76] 224 yes/no
Dietary intake and micronutrient concentrations
Serum calcium (mg/dL) - CAD [23] 4 unit
Omega-6 (mg/dL) - CAD [49] 11 sd
Inflammatory and thrombotic biomarkers
Colony stimulating factor 1 levels (ng/mL) - CAD [30] 7 sd
Interleukin-6 receptor (ng/mL) - CAD [61] 17 sd
Protein C (mg/dL) - CAD [20] 6 unit
Lifestyle, education and behavior
Intelligence (score) - CAD [80] 121 sd
Smoking Initiation - CAD [53] 342 yes/no
Lipids and lipoproteins
Apolipoprotein B (mg/dL) - CAD [52] 12 sd
Linoleic acid (mg/dL) - MI [69] 3 unit
High density lipoprotein cholesterol (mg/dL) - CAD [54] 202 unit
Low density lipoprotein cholesterol (mg/dL) - CAD [54] 164 unit
Total cholesterol (mg/dL) - CAD [49] 11 sd
Triglycerides (mg/dL) - CAD [17] 140 sd
ApoE*ε3 (inversed-normalized change in each apoE isoform) - CAD [67] 5 sd
Other diseases and traits
Lymphocyte count - CAD [13] 289 sd
Childhood Obesity - CAD [35] 9 yes/no
Depression - CAD [67] 102 yes/no
Educational attainment - CAD [9] 1271 sd
Mean arterial pressure (mmHg) - CAD [46] 124 sd
Obesity - CAD [18] 97 yes/no
Visceral adiposity - CAD [77] 221 yes/no
Other metabolites/biomarkers
Serum urate (mg/dL) - CAD [81] 98 unit
Steroids and hormones
Sex hormone binding globulin (nmol/L) - CAD [72] 357 sd
0.25
0.50
1.0
2.0
4.0
OR (95%CI) per 1 unit increase in exposure

Figure 4 B


IVW WM Egger
Exposure (unit) - Outcome [Ref]
No of SNPs
Contrast
Anthropometrics
Body Mass Index (kg/m2) - Stroke [64] 90 sd
Height (cm) - Stroke [78] 2084 sd
Maternal effect on birth weight (g) - LVIS [56] 256 sd
Waist-hip ratio - Stroke [34] 324 sd
Clinical measurements
Diastolic blood pressure (mmHg) - Stroke [45] 462 sd
Systolic blood pressure (mmHg) - Stroke [45] 460 sd
Diabetes and related biomarkers
Glucose (mg/dL) - IS [59] 10 unit
Hemoglobin A1c (%) - LVIS [63] 333 unit
Dietary intake and micronutrient concentrations
Iron (mg/dL) - CIS [27] 11 sd
Transferrin (mg/dL) - CIS [27] 11 sd
Transferrin saturation (%) - CIS [27] 11 sd
Urinary Sodium (mmol/L) - IS [16] 51 sd
Vitamin K (nmol/L) - LVIS [28] 4 unit
Inflammatory and thrombotic biomarkers
Factor XI (unit/dl) - IS [25] 26 unit
Factor VII (unit/dl) - IS [35] 1637 unit
Interleukin-6 receptor (ng/mL) - Stroke [61] 17 sd
Lipids and lipoproteins
Apolipoprotein B (mg/dL) - IS [58] 213 sd
Low density lipoprotein cholesterol (mg/dL) - Stroke [62] 11 unit
Linoleic acid (mg/dL) - LVIS [57] 71 unit
Other diseases and traits
Atrial Fibrillation - Stroke [75] 116 yes/no
Educational attainment - Stroke [9] 1271 sd
Platelet Count - IS [26] 219 sd
0.25
0.50
1.0
2.0
4.0
OR (95%CI) per 1 unit increase in exposure