



1 Brief Report

2 Associations of Circulating Biomarkers with Disease Risks: a 3 Two-Sample Mendelian Randomization Study

4 Abdulkadir Elmas^{1,†}, Kevin Spehar^{2,3,4,5,†}, Ron Do¹, Joseph M. Castellano^{2,3,4,5}, Kuan-lin Huang^{1,*}

5 ¹ Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY
6 10029, USA

7 ² Nash Family Department of Neuroscience, Friedman Brain Institute, Icahn School of Medicine at Mount Si-
8 nai, New York, NY, USA

9 ³ Ronald M. Loeb Center for Alzheimer's Disease, Icahn School of Medicine at Mount Sinai, New York, NY,
10 USA

11 ⁴ Department of Neurology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

12 ⁵ Black Family Stem Cell Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

13 † These authors contributed equally to this work

14 * Correspondence: kuan-lin.huang@mssm.edu

15

16 **Abstract: Background:** Circulating biomarkers play a pivotal role in personalized medicine, offer-
17 ing potential for disease screening, prevention, and treatment. Despite established associations
18 between numerous biomarkers and diseases, elucidating their causal relationships is challenging.
19 Mendelian Randomization (MR) can address this issue by employing genetic instruments to dis-
20 cern causal links. Additionally, using multiple MR methods with overlapping results enhances the
21 reliability of discovered relationships. **Methods:** Here we report an MR study using multiple
22 methods, including inverse variance weighted, simple mode, weighted mode, weighted median,
23 and MR Egger. We use the MR-base resource (v0.5.6)¹ to evaluate causal relationships between 212
24 circulating biomarkers (curated from UK Biobank analyses by Neale lab and from Shin et al. 2014,
25 Roederer et al. 2015, and Kettunen et al. 2016)²⁻⁴ and 99 complex diseases (curated from several
26 consortia by MRC IEU and Biobank Japan). **Results:** We report novel causal relationships found by
27 4 or more MR methods between glucose and bipolar disorder (Mean Effect Size estimate across
28 methods: 0.39) and between cystatin C and bipolar disorder (Mean Effect Size: -0.31). Based on
29 agreement in 4 or more methods, we also identify previously known links between urate with gout
30 and creatine with chronic kidney disease, as well as biomarkers that may be causal of cardiovas-
31 cular conditions: apolipoprotein B, cholesterol, LDL, lipoprotein A, and triglycerides in coronary
32 heart disease, as well as lipoprotein A, LDL, cholesterol, and apolipoprotein B in myocardial in-
33 farction. **Conclusions:** This Mendelian Randomization study not only corroborates known causal
34 relationships between circulating biomarkers and diseases but also uncovers two novel biomarkers
35 associated with bipolar disorder that warrant further investigation. Our findings provide insight
36 into understanding how biological processes reflecting circulating biomarkers and their associated
37 effects may contribute to disease etiology, which can eventually help improve precision diagnostics
38 and intervention.

39 **Keywords:** Biomarkers; Metabolome; Proteome; Mendelian Randomization; Human Disease

40

41 1. Introduction

42 Molecular abnormalities detected in the blood that cause complex diseases represent
43 an opportunity to identify biomarkers for both preventive and therapeutic interventions.
44 These circulating biomolecules include proteins (enzymes and hormones), lipids, and
45 metabolites that reflect physiological states of organ functions, immune response, and

46 metabolism. Despite substantial observational evidence linking systemic biomarker lev-
47 els with diverse health conditions, their causal relationships to complex diseases remain
48 to be established, especially for diseases for which no reliable biomarkers exist. Previous
49 studies that identify correlations often do not delineate cause and effect, thereby limiting
50 the translational value of the findings. For example, while an increased incidence of im-
51 paired glucose metabolism has been demonstrated in patients with bipolar disorder
52 across multiple studies, a causal relationship between the two has yet to be established.⁵⁻⁷

53 The application of Mendelian Randomization (MR) represents an approach in ad-
54 dressing this critical gap. By utilizing genetic instruments as proxies for biomarker levels,
55 MR, under specific assumptions, can control for confounding factors and reverse causa-
56 tion, offering insights into the causal effects of biomarkers on disease risks. Lipids have
57 been particularly well studied in MR, where previous studies have demonstrated the
58 causal relationships between LDL-c and coronary artery disease, and HDL-c and breast
59 cancer.⁸⁻¹¹ Other studies have also identified likely causal relationships between
60 homocysteine and stroke,¹² metabolic syndrome,¹³ as well as tyrosine and Type 2 Dia-
61 betes.¹⁴ However, the effects of many other biomarkers remain unclear, and existing MR
62 studies have often been limited to specific biomarkers or disease categories.

63 The power of MR has been boosted by the availability of large genome-wide asso-
64 ciation studies (GWAS), published in more than 379 studies including those conducted
65 using the UK Biobank dataset that has recently released a resource of plasma biomarker
66 data measured by nuclear magnetic resonance (NMR) in addition to other lab markers.¹⁵
67 Herein, we conducted a comprehensive MR analysis encompassing a broad spectrum of
68 212 biomarkers—including 115 circulating biomolecules measured through NMR⁴—and
69 99 human diseases to unveil previously obscured relationships of the circulating bi-
70 omarkers' role in disease etiology that may be leveraged for personalized prevention.

71 2. Results

72 Based on the final 212 exposures (circulating biomarkers) and 99 outcomes (diseas-
73 es) from the MRC IEU OpenGWAS database¹⁶ (Methods, STROBE-MR checklist), we
74 analyzed exposure-outcome relationships using 5 different MR analysis methods: inverse
75 variance weighted, MR Egger, simple mode, weighted median, and weighted mode. To
76 ensure our results were robust, we focused on findings that had a significant (Bonferroni
77 p-value < 0.05) association by MR-Egger and matched in 4 or more methods (with same
78 effect direction and raw p-value < 0.05) based on MR-Egger's robustness to directional
79 pleiotropy relevant to circulating biomarkers. We identified a total 21 significant bi-
80 omarker exposure vs. disease outcome associations (Figure 1). A list of all significant re-
81 lations, including IEU GWAS study IDs for each exposure/outcomes, and the summary
82 statistics from the MR analyses are included in Supplementary Table 1.

83 We provide validation for our analysis by demonstrating agreement using 4 or more
84 analysis methods that have been previously demonstrated in MR studies such as urate
85 with gout¹⁷ and LDL,¹⁸ triglycerides,¹⁰ lipoprotein A,¹⁹ and apolipoprotein B²⁰ with coro-
86 nary heart disease. We also found previously demonstrated relationships shown in MR
87 studies between lipoprotein A²¹ and LDL²² with myocardial infarction. Additionally, we
88 did not find any relationships that showed conflicting directionality to the previous lit-
89 erature. While many identified relationships are only maintained in a few methods, pre-
90 viously supported relationships show strong alignment in multiple MR analysis meth-
91 ods, demonstrating the robustness of associations found using multiple approaches.

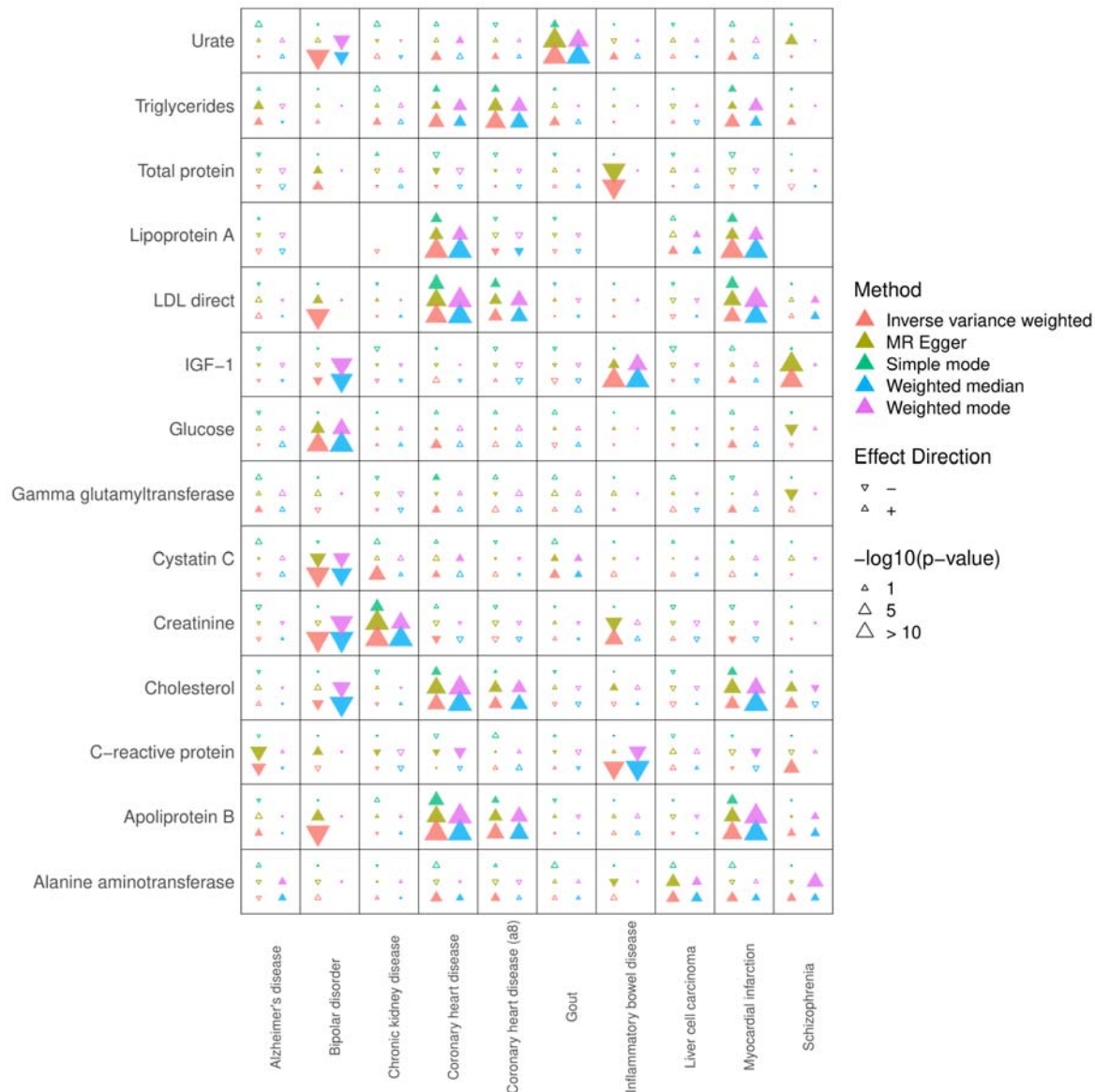
92 For cardiometabolic traits, we demonstrated a causal relationship between total
93 cholesterol and coronary heart disease, and this association has previously been shown in
94 a meta-analysis and systematic review.²³ We also showed a causal relationship between
95 total cholesterol and myocardial infarction as well as apolipoprotein B and myocardial
96 infarction, both of which has been previously demonstrated by non-MR studies.^{24,25}

97 We also found a causal relationship between serum creatinine and chronic kidney
98 disease (CKD). However, serum creatinine is an important biomarker for measuring

99
100
101
102
103
104
105
106
107
108
109
110

kidney function, as it is used to clinically estimate glomerular filtration rate to help diagnose CKD.^{26,27} While this connection could explain this result, there may be creatinine-specific effects that may cause or exacerbate CKD which remain to be characterized. We also identified a trend for alanine transferase with liver cell carcinoma that agreed for 4+ methods when using raw p-value cut-offs of < 0.05.

Of note, we discovered a strong, direct relationship between glucose and bipolar disorder and a strong, inverse relationship between cystatin C levels and bipolar disorder which, to the best of our knowledge, has never been directly reported before. Previous studies have demonstrated an association between impaired glucose metabolism and bipolar disorder in addition to increased prevalence of pre-diabetes and type 2 diabetes mellitus.^{7,28,29} Table 1 lists the identified exposure-outcome relationships along with previous supporting literature and evidence type.



111
112
113
114
115

Figure 1. Heatmap summarizing the key findings of this study, including the most significant findings from the MR Egger analysis. MR analyses were conducted using 5 methods: Inverse variance weighted, MR Egger, Simple mode, Weighted Median, and Weighted Mode. Larger arrow sizes correspond to more significant results and the directionality of the arrow indicates a posi-

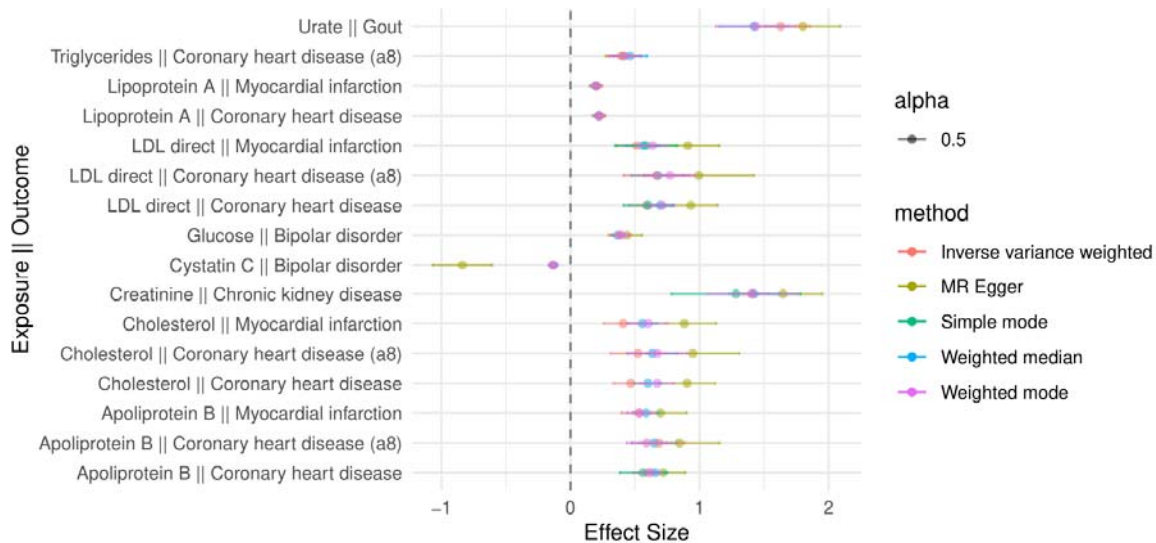
116 tively-correlated relationship, i.e., higher biomarkers associated with increased risk (upward ar-
 117 row) or inverse (downward arrow) relationship. We displayed all exposures and outcomes that
 118 demonstrated at least one significant (Bonferroni p-value < 0.05) relationship by at least one MR
 119 method.

120 **Table 1.** Significant exposure-outcomes relationships consistently identified across 4 or more MR
 121 methods. Previous literature demonstrating these relationships and the type of evidence support-
 122 ing each relationship are also detailed. In evidence type, retrospective, prospective, and
 123 cross-sectional indicates cohort studies that demonstrate associations of the biomarker with the
 124 disease outcome, whereas meta-analysis and systematic review represent syntheses of such studies
 125 that do not suggest causality.

| Exposure | Outcome | Previous Evidence | Evidence Type |
|------------------|------------------------|-------------------|---|
| LDL | Coronary Heart Disease | 18 | MR |
| Triglycerides | Coronary Heart Disease | 10 | MR |
| Lipoprotein A | Coronary Heart Disease | 19 | MR |
| Apolipoprotein B | Coronary Heart Disease | 20 | MR |
| Cholesterol | Coronary Heart Disease | 23,30 | Meta-Analysis, Systematic Review, MR |
| Urate | Gout | 17 | MR |
| Lipoprotein A | Myocardial Infarction | 21 | MR |
| LDL | Myocardial Infarction | 22 | MR |
| Cholesterol | Myocardial Infarction | 24 | Retrospective |
| Apolipoprotein B | Myocardial Infarction | 25,31 | Prospective, MR |
| Serum Creatinine | Chronic Kidney Disease | 26,27 | MR, Meta-Analysis |
| Glucose | Bipolar Disorder | 28,29 | Cross Sectional, Meta-Analysis, Systematic Review |
| Cystatin C | Bipolar Disorder | N/A | N/A |

126 Figure 2 illustrates the top-16 significant (Bonferroni p-value < 0.05) biomarker ex-
 127 posure-disease outcome associations with agreement in effect directionality in 4 or more
 128 methods (effect size > |0.1|).

129 Several biomarkers are associated with an increased risk of disease. The average ef-
 130 fect size of total cholesterol and coronary heart disease is 0.66, ranging from 0.90 ± 0.22
 131 (MR Egger) to 0.47 ± 0.13 (IVW). The association of total cholesterol and myocardial in-
 132 farction has the average effect size of 0.61 with a range of (0.26, 1.13). For the
 133 apolipoprotein B and myocardial infarction, the average effect size is estimated as 0.59
 134 with a range of (0.39, 0.90). For the glucose on bipolar disorder the average effect estimate
 135 is 0.39 with a range of (0.29, 0.55), while cystatin C has an inverse relationship on bipolar
 136 disorder with the average effect size of -0.31 ranging from -0.107 to -1 (Supplementary
 137 Table 1).
 138



139
140
141
142
143
144
145
146
147

Figure 2. Forest Plot of Mendelian Randomization Analysis Results Assessing the Causal Effects of Circulating Biomarkers on Diseases. The forest plot illustrates the estimated causal effects of top-16 significantly associated (Bonferroni p-value < 0.05) biomarkers on the risk of various diseases based on Mendelian randomization (MR) analysis consistent across 4 or more methods. Each row represents a different MR association, with the biomarker as the exposure and the disease as the outcome. The points represent different effect size estimates from different MR methods and the 95% confidence intervals (based on standard errors of the effect size estimates) are displayed by the horizontal lines forming around the average effect sizes of different methods.

148
149

3. Methods

3.1. Datasets

150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173

This study is a two-sample MR study based on publicly available GWAS summary statistics data from MR-base resource (v0.5.6).¹ The study design follows the Strengthening the Reporting of Observational Studies in Epidemiology using Mendelian Randomisation (STROBE-MR) checklist, which is included in Supplementary information.

The MR-base platform, developed by the MRC Integrative Epidemiology Unit at the University of Bristol, serves both as a database and an analytical framework accessed through the TwoSampleMR R package (v0.5.7). MR-base offers access to multiple GWAS data including the MRC Integrative Epidemiology Unit (IEU) OpenGWAS database consisting of 26k GWAS reported in at least 379 different studies, covering ~40k individuals (in median) per study.^{1,16} Based on MRbase, we utilized all accessible exposure traits grouped under "Metabolites" category as well as other circulating biomarkers, and all available outcomes categorized as "Disease". All diseases we used are from datasets with IDs beginning with "ieu" (GWAS summary datasets generated by many consortia curated by IEU) and "bbj" (Biobank Japan), and the biomarkers are from "ukb" (UK Biobank analyses by IEU or Neale lab covering 28 biomarkers) and "met" (human blood biomarkers, immune markers, and circulating biomarkers analyzed by Shin et al. 2014 covering 76 biomarkers and Kettunen et al. 2016 covering 115 biomarkers).^{2,4} In instances where there are multiple summary datasets for the same trait (i.e., from different GWAS cohorts), we selected the top one from them based on the largest "sample size," "year," "number of SNPs," or "number of cases and controls," prioritizing the information available in this specified order. A list of all biomarkers used as exposures and information for their corresponding GWAS studies are available in Supplementary Table 2. Overall, we retained a total of 212 exposures (biomarkers) and 99 outcomes (diseases) for MR analyses.

174 Given that this is a two-sample MR study, we provide two major justifications: (1)
175 Similarity of the Genetic Variant-Exposure Associations Between the Exposure and
176 Outcome Samples: Genetic instruments for the exposure were primarily derived from
177 GWAS of circulating biomarkers conducted using data from the UK Biobank in addition
178 to Shin and Kettunen et al., all of which predominantly consists of individuals of Euro-
179 pean ancestry. The outcome data were mainly sourced from case-control GWAS of Eu-
180 ropean populations that likely share genetic architectures similar to those of the exposure
181 study participants. (2) Number of Individuals Who Overlap Between the Exposure and
182 Outcome Studies: Given the separate sources of the genetic data for exposures and out-
183 comes (various case-control studies), and the timing of data availability and publication,
184 it is reasonable to assume minimal overlap in the individual participants across these
185 studies. Particularly, all significant results were found based on exposure (biomarker)
186 GWAS from UK Biobank, which has its first genetic data published in 2018. The signifi-
187 cant outcome GWAS for bipolar disorder, cardiovascular disease (CVD), myocardial in-
188 farction (MI), gout, and chronic kidney disease, all predated 2018. This temporal dis-
189 crepancy between the data collection phases further supports the assumption of negli-
190 gible overlap bias in our MR estimates.

191 3.2. MR Methods

192 The three core assumptions for Mendelian Randomization (MR) analysis are:

193 Relevance: The genetic instruments (e.g., single nucleotide polymorphisms, SNPs)
194 used as instrumental variables (IVs) should be associated with the exposure (circulating
195 biomarkers in this case). This assumption ensures that the IVs are strong enough to in-
196 fluence the exposure. Herein, 2685 SNPs were found associated with 198 circulating bi-
197 omarker exposures subject to the IV assumptions that were imposed by the standard
198 two-sample MR approach (association threshold, $P < 5 \times 10^{-8}$; LD clumping cutoff $r^2 > 0.001$
199 within 10Mb window) and were used for further analyses (Supplementary Table 3).

200 Independence: The genetic instruments should be independent of any confounding
201 factors that may affect both the exposure and the outcome (disease). This assumption is
202 justified by the random allocation of genetic variants. Further, each of the GWAS studies
203 carefully control for covariates that may confound genetic associations. For example, in
204 the UK Biobank biomarker GWAS conducted by the Neale lab, covariates include age,
205 sex, age², age*sex, age*sex², and the first 20 PCs, making the final genetic instruments
206 less prone to confounding. Additionally, the prior knowledge of demographic con-
207 sistency can likely minimize the impact of unmeasured confounders due to similar de-
208 mographic and genetic backgrounds.

209 Exclusion restriction: The genetic instruments should affect the outcome (disease)
210 only through their effects on the exposure (circulating biomarkers) and not through any
211 other direct or indirect pathways. This assumption ensures that the IVs influence the
212 outcome solely via the exposure of interest. Critically, MR Egger was employed as a
213 primary MR method in this study to detect and adjust for pleiotropy, where genetic
214 variants may influence the outcome via pathways other than through the exposure.

215 To enhance robustness, we inferred causality using multiple MR methods, specifi-
216 cally inverse variance weighted, simple mode, weighted mode, weighted median, and
217 MR Egger. The inverse variance weighted (IVW) method conducts regression analysis
218 between the SNP-exposure effect size and the SNP-outcome effect size, giving more im-
219 portance to SNPs with the lowest standard error in SNP-outcome association. The IVW
220 method relies on either all variants being valid instruments or the presence of balanced
221 horizontal pleiotropy, where the combined horizontal pleiotropic effects of individual
222 instruments cancel out. Additionally, it assumes that such pleiotropic effects are not de-
223 pendent on instrument strength across all variants, a concept known as the Instrument
224 Strength Independent of Direct Effects (InSIDE) assumption.³² The simple mode method
225 operates under the assumption that the most common value (i.e., mode) for horizontal
226 pleiotropy is zero, known as the ZERO Modal Pleiotropy Assumption (ZEMPA), regard-

less of the specific type of horizontal pleiotropy.³³ Likewise, the weighted MBE method operates under the supposition that the most substantial weights in the k subsets originate from valid instruments, adhering to ZEMPA assumption, irrespective of the particular form of horizontal pleiotropy. The weighted median method relies on the assumption that over half of the weight comes from reliable instruments, irrespective of the nature of horizontal pleiotropy.³⁴ MR-Egger permits all genetic variants to exhibit pleiotropic effects, but it necessitates that these effects are independent of the SNP–exposure associations, following the InSIDE principle.³⁵ Notably, the method can correct for the average horizontal pleiotropy observed across all variants considered in the study. Considering the recognized pleiotropic effects commonly seen in biomarkers and their relevance to health and disease,³⁶ we used MR-Egger as the primary method for determining important exposure-outcome associations in our analyses and employed other sensitivity methods for validation.

For each method, the associations are reported with effect size and corresponding p -value statistics. The raw p -values are multi-testing corrected using the Bonferroni method. The confidence intervals (95%) are also calculated from standard errors of the test statistics. Overall, robustness of the results is ensured based on the use of MR-Egger as discovery method, agreement across at least 4 MR methods, and Bonferroni correction.

4. Discussion

We analyzed exposure-outcome relationships using 5 MR analysis methods to identify causal relationships between circulating biomarkers and diseases. In this study, significant results are defined as agreement between 4 or more analysis methods. Many of our demonstrated results replicated results from previous MR studies and other literature and none of our findings run contrary to previous literature. Additionally, our analysis also discovered two relationships of note: a direct relationship between glucose and bipolar disorder and an inverse relationship between cystatin C and bipolar disorder.

Bipolar disorder is a neuropsychiatric disorder with multifactorial causes (genetic, trauma, exposure to certain medications, etc.) Our findings demonstrate two potential circulating markers that may contribute to development of bipolar disorder, high glucose and low cystatin C. In our study, the effect size for glucose and bipolar disorder is 0.39 and for cystatin C and bipolar disorder is -0.31 which is comparable to the average effect size for triglycerides and coronary heart disease (a_8) of 0.42 and greater than those of triglycerides and myocardial infarction (0.20), lipoprotein A and coronary heart disease (0.22), and lipoprotein A and myocardial infarction (0.21) (Supplementary Table 1). These results suggest that these relationships are likely to be of clinical significance and warrants further study. Although high cystatin C levels have been linked with major depressive disorder,³⁷ no studies, to the best of our knowledge, have demonstrated an association between cystatin C and bipolar disorder. Levels of Cystatin C, a natural inhibitor of cysteine proteases, are typically used clinically to assess kidney function. A common treatment for bipolar disorder is lithium, which has been shown to decrease renal function and can lead to elevated cystatin C levels,³⁸ which may lend further insight to the mechanism behind lithium's use as a treatment, however this remains to be confirmed.

A connection between impaired glucose metabolism and bipolar disorder has been well documented in the literature, with over half of patients diagnosed with bipolar disorder also having insulin resistance, impaired glucose tolerance, or type 2 diabetes.⁷ Furthermore, it has been shown that modulating the PI3K/Akt insulin signaling pathway may be a mechanism of lithium for the treatment of bipolar disorder.³⁹ The PI3K/Akt pathway helps regulate glucose metabolism in the hippocampus, cerebellum, and olfactory bulb⁴⁰ which are vulnerable sites of gray matter loss in bipolar disorder.⁴¹ It has been speculated that long-term disruption of this pathway can lead to mitochondrial dysfunction and energy dysregulation which may affect brain processes contributing to bipolar

280 disorder.³⁹ Therefore, our identification of a causal relationship between high glucose
281 levels and bipolar disorder further supports the current body of evidence, and our find-
282 ings for cystatin C highlight a novel finding that warrants investigation in subsequent
283 studies. Our findings suggest that future studies aiming to reduce the risk of bipolar
284 disorder could explore the lifestyle and clinical interventions that reduce blood glucose
285 and maintain kidney function.

286 Although our study focused on links that were aligned with 4 or more MR methods,
287 other notable relationships were identified that were aligned with 2-3 MR methods
288 (Figure 1). For example, C-reactive protein (CRP) and Alzheimer's disease were causally
289 linked in 2 methods (MR Egger and IVW) which has been previously noted in previous
290 non-MR studies.⁴²

291 A limitation of this study is that much of the exposure GWAS data comes from the
292 UK biobank, a large-scale volunteer databank, which has a biased representation of Eu-
293 ropeans and results may not generalize outside UK.^{43,44} However, UK biobank provides
294 the largest GWAS datasets of most biomarkers studied herein and adequate statistical
295 power; this bias may be overcome as other biobanks with diverse populations continue to
296 expand. MR studies have their own limitations, including but not limited to: potential
297 confounding factors, limitations of estimating associations for binary outcomes, plei-
298 otropic effects which should be mitigated by our use of MR Egger, and population strat-
299 ification.⁴⁵⁻⁴⁷

300 Overall, by utilizing several MR methods as cross-validation, our study successfully
301 reaffirmed several established connections in cardiovascular diseases, gout, and kidney
302 diseases, while uncovering intriguing novel biomarkers that may be causal for bipolar
303 disorder. These findings highlight the underappreciated roles that circulating biomarkers
304 may play in disease mechanisms, which should open new avenues for targeted research
305 and motivate development of precise diagnostic tools and therapeutic interventions.
306

307 **Supplementary Materials:** The following supporting information can be downloaded at:
308 www.mdpi.com/xxx/s1, Figure S1: title; Table S1: title; Video S1: title.

309 **Author Contributions:** AE designed and conducted the analyses. KS performed the literature
310 search and interpreted the results. KS, AE, JMC, and KH wrote the manuscript. KH supervised the
311 study. All authors read, edited, and approved the manuscript.

312 **Funding:** This research was funded by National Institutes of Health, grant number:
313 1R35GM138113-03P1.

314 **Institutional Review Board Statement:** Not applicable. Our analyses were conducted using pub-
315 licly available summary statistics data from GWAS, not individual-level data; therefore, this is a
316 non-human subject research that requires no additional ethical approval.

317 **Data Availability Statement:** Data are contained within the article and supplementary materials.

318 **Data Availability:** GWAS summary statistics data used can be found on MR-base resource (v0.5.6)
319 through the ieu open GWAS project at the URL: <https://gwas.mrcieu.ac.uk/>.

320 **Code Availability:** The source code for the Mendelian randomization analyses is available at
321 github.com/Huang-lab/MR.

322 **Acknowledgments:** The authors wish to acknowledge providers of GWAS summary datasets. The
323 authors thank all members of the Huang lab for constructive discussion. Large language models
324 (LLM) may have been used in the initial drafts of coding and writing of this work. All final codes
325 and texts have been extensively edited and verified by the authors. This work was supported in
326 part through the computational and data resources and staff expertise provided by Scientific
327 Computing and Data at the Icahn School of Medicine at Mount Sinai and supported by the Clinical
328 and Translational Science Awards (CTSA) grant UL1TR004419 from the National Center for Ad-
329 vancing Translational Sciences. Research reported in this publication was also supported by the
330 Office of Research Infrastructure of the National Institutes of Health under award number
331 S10OD026880 and S10OD030463. The content is solely the responsibility of the authors and does

332 not necessarily represent the official views of the National Institutes of Health. This work was
333 supported by NIH NIGMS R35GM138113 and ACS RSG-22-115-01-DMC to KH and RF1AG072300
334 to JMC.

335 **Conflicts of Interest:** The authors declare no conflict of interest.

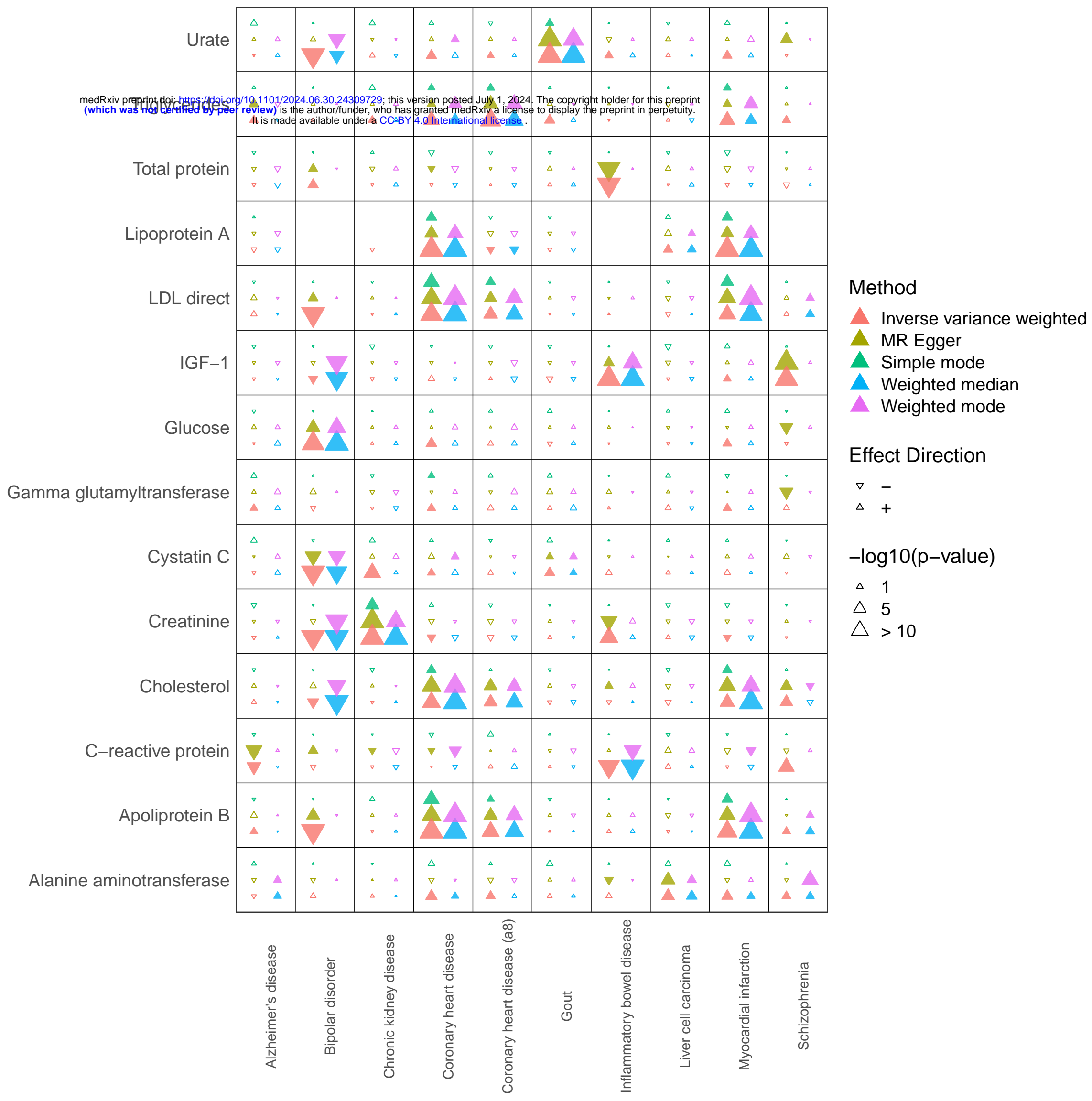
336 References

- 337 1. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human
338 phenome. *Elife* 2018;7. DOI: 10.7554/eLife.34408.
- 339 2. Shin SY, Fauman EB, Petersen AK, et al. An atlas of genetic influences on human blood metabolites. *Nat Genet*
340 2014;46(6):543-550. DOI: 10.1038/ng.2982.
- 341 3. Roederer M, Quaye L, Mangino M, et al. The genetic architecture of the human immune system: a bioresource for
342 autoimmunity and disease pathogenesis. *Cell* 2015;161(2):387-403. DOI: 10.1016/j.cell.2015.02.046.
- 343 4. Kettunen J, Demirkan A, Wurtz P, et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel
344 systemic effects of LPA. *Nat Commun* 2016;7:11122. DOI: 10.1038/ncomms11122.
- 345 5. Ruzickova MS, C.
346 Garnham, J., Alda M. Clinical Features of Bipolar Disorder With and Without Comorbid Diabetes Mellitus. *The Canadian Journal of*
347 *Psychiatry* 2003;48(7):431-502.
- 348 6. Calkin CV, Ruzickova M, Uher R, et al. Insulin resistance and outcome in bipolar disorder. *Br J Psychiatry* 2015;206(1):52-7.
349 DOI: 10.1192/bjp.bp.114.152850.
- 350 7. Lojko D, Owecki M, Suwalska A. Impaired Glucose Metabolism in Bipolar Patients: The Role of Psychiatrists in Its
351 Detection and Management. *Int J Environ Res Public Health* 2019;16(7). DOI: 10.3390/ijerph16071132.
- 352 8. Beeghly-Fadiel A, Khankari NK, Delahanty RJ, et al. A Mendelian randomization analysis of circulating lipid traits and
353 breast cancer risk. *Int J Epidemiol* 2020;49(4):1117-1131. DOI: 10.1093/ije/dyz242.
- 354 9. Burgess S, Harshfield E. Mendelian randomization to assess causal effects of blood lipids on coronary heart disease:
355 lessons from the past and applications to the future. *Curr Opin Endocrinol Diabetes Obes* 2016;23(2):124-30. DOI:
356 10.1097/MED.000000000000230.
- 357 10. Holmes MV, Asselbergs FW, Palmer TM, et al. Mendelian randomization of blood lipids for coronary heart disease. *Eur*
358 *Heart J* 2015;36(9):539-50. DOI: 10.1093/eurheartj/ehv571.
- 359 11. Thomas DG, Wei Y, Tall AR. Lipid and metabolic syndrome traits in coronary artery disease: a Mendelian randomization
360 study. *J Lipid Res* 2021;62:100044. DOI: 10.1194/jlr.P120001000.
- 361 12. Larsson SC, Traylor M, Markus HS. Homocysteine and small vessel stroke: A mendelian randomization analysis. *Ann*
362 *Neurol* 2019;85(4):495-501. DOI: 10.1002/ana.25440.
- 363 13. Lee HS, In S, Park T. The Homocysteine and Metabolic Syndrome: A Mendelian Randomization Study. *Nutrients*
364 2021;13(7). DOI: 10.3390/nu13072440.
- 365 14. Jager S, Cuadrat R, Wittenbecher C, et al. Mendelian Randomization Study on Amino Acid Metabolism Suggests Tyrosine
366 as Causal Trait for Type 2 Diabetes. *Nutrients* 2020;12(12). DOI: 10.3390/nu12123890.
- 367 15. Julkunen H, Cichonska A, Tiainen M, et al. Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals
368 from the UK Biobank. *Nat Commun* 2023;14(1):604. DOI: 10.1038/s41467-023-36231-7.
- 369 16. Elsworth B, Lyon M, Alexander T, et al. The MRC IEU OpenGWAS data infrastructure. *bioRxiv* 2020. DOI:
370 10.1101/2020.08.10.244293.
- 371 17. Adams CD, Boutwell BB. Using multiple Mendelian randomization approaches and genetic correlations to understand
372 obesity, urate, and gout. *Sci Rep* 2021;11(1):17799. DOI: 10.1038/s41598-021-97410-4.

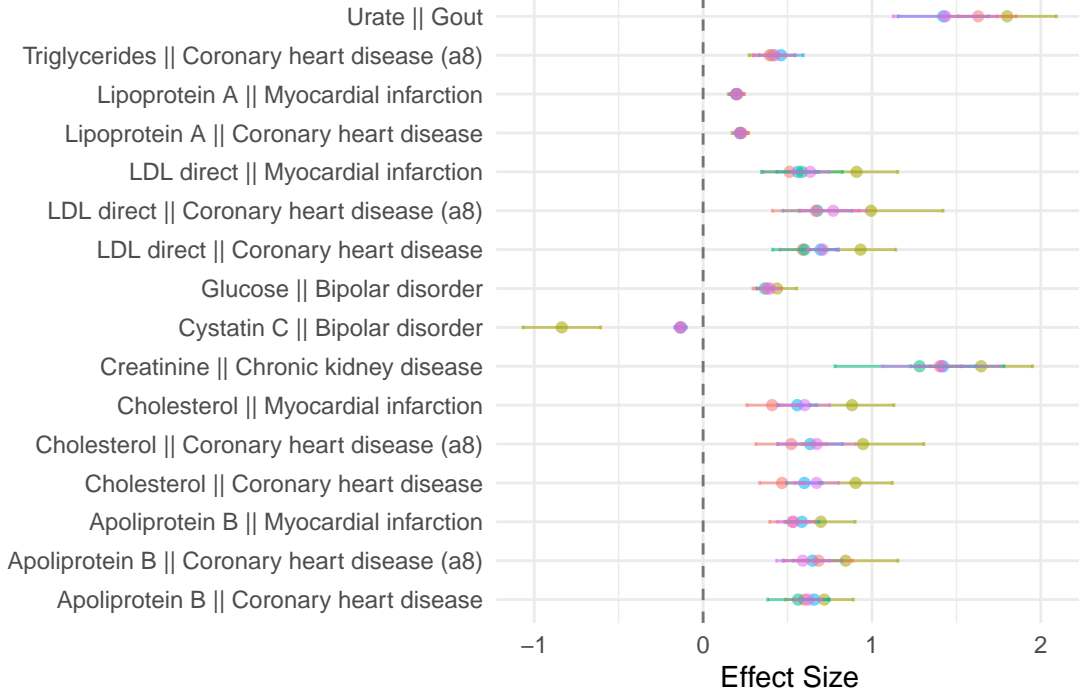
- 373 18. Ference BA. Mendelian randomization studies: using naturally randomized genetic data to fill evidence gaps. *Curr Opin*
374 *Lipidol* 2015;26(6):566-71. DOI: 10.1097/MOL.0000000000000247.
- 375 19. Wang S, Zha L, Chen J, et al. The relationship between lipoprotein(a) and risk of cardiovascular disease: a Mendelian
376 randomization analysis. *Eur J Med Res* 2022;27(1):211. DOI: 10.1186/s40001-022-00825-6.
- 377 20. Richardson TG, Sanderson E, Palmer TM, et al. Evaluating the relationship between circulating lipoprotein lipids and
378 apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. *PLoS Med*
379 2020;17(3):e1003062. DOI: 10.1371/journal.pmed.1003062.
- 380 21. Nordestgaard BG, Langsted A. Lipoprotein (a) as a cause of cardiovascular disease: insights from epidemiology, genetics,
381 and biology. *J Lipid Res* 2016;57(11):1953-1975. DOI: 10.1194/jlr.R071233.
- 382 22. Pham K, Mulugeta A, Lumsden A, Hyppönen E. Genetically instrumented LDL-cholesterol lowering and multiple disease
383 outcomes: A Mendelian randomization phenome-wide association study in the UK Biobank. *Br J Clin Pharmacol*
384 2023;89(10):2992-3004. DOI: 10.1111/bcp.15793.
- 385 23. Peters SA, Singhathe Y, Mackay D, Huxley RR, Woodward M. Total cholesterol as a risk factor for coronary heart disease
386 and stroke in women compared with men: A systematic review and meta-analysis. *Atherosclerosis* 2016;248:123-31. DOI:
387 10.1016/j.atherosclerosis.2016.03.016.
- 388 24. Kim MK, Han K, Kim HS, et al. Cholesterol variability and the risk of mortality, myocardial infarction, and stroke: a
389 nationwide population-based study. *Eur Heart J* 2017;38(48):3560-3566. DOI: 10.1093/eurheartj/ehx585.
- 390 25. Walldius G, Jungner I, Holme I, Aastveit AH, Kolar W, Steiner E. High apolipoprotein B, low apolipoprotein A-I, and
391 improvement in the prediction of fatal myocardial infarction (AMORIS study): a prospective study. *Lancet*
392 2001;358(9298):2026-33. DOI: 10.1016/S0140-6736(01)07098-2.
- 393 26. Park S, Lee S, Kim Y, et al. Mendelian randomization reveals causal effects of kidney function on various biochemical
394 parameters. *Commun Biol* 2022;5(1):713. DOI: 10.1038/s42003-022-03659-4.
- 395 27. Levey AS, de Jong PE, Coresh J, et al. The definition, classification, and prognosis of chronic kidney disease: a KDIGO
396 Controversies Conference report. *Kidney Int* 2011;80(1):17-28. DOI: 10.1038/ki.2010.483.
- 397 28. Mansur RB, Rizzo LB, Santos CM, et al. Impaired glucose metabolism moderates the course of illness in bipolar disorder. *J*
398 *Affect Disord* 2016;195:57-62. DOI: 10.1016/j.jad.2016.02.002.
- 399 29. Vancampfort D, Mitchell AJ, De Hert M, et al. Prevalence and predictors of type 2 diabetes mellitus in people with bipolar
400 disorder: a systematic review and meta-analysis. *J Clin Psychiatry* 2015;76(11):1490-9. DOI: 10.4088/JCP.14r09635.
- 401 30. Liao LZ, Li WD, Liu Y, Li JP, Zhuang XD, Liao XX. Exploring the causal pathway from omega-6 levels to coronary heart
402 disease: A network Mendelian randomization study. *Nutr Metab Cardiovasc Dis* 2020;30(2):233-240. DOI:
403 10.1016/j.numecd.2019.09.013.
- 404 31. Ding Y, Yang S, He M, Fan S, Tao X, Lu W. Lipid Metabolism Traits Mediate the Effect of Psoriasis on Myocardial
405 Infarction Risk: A Two-Step Mendelian Randomization Study. *Metabolites* 2023;13(9). DOI: 10.3390/metabo13090976.
- 406 32. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using
407 summarized data. *Genet Epidemiol* 2013;37(7):658-65. DOI: 10.1002/gepi.21758.
- 408 33. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal
409 pleiotropy assumption. *Int J Epidemiol* 2017;46(6):1985-1998. DOI: 10.1093/ije/dyx102.
- 410 34. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid
411 Instruments Using a Weighted Median Estimator. *Genet Epidemiol* 2016;40(4):304-14. DOI: 10.1002/gepi.21965.
- 412 35. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias
413 detection through Egger regression. *Int J Epidemiol* 2015;44(2):512-25. DOI: 10.1093/ije/dyv080.

- 414 36. Smith CJ, Sinnott-Armstrong N, Cichonska A, et al. Integrative analysis of metabolite GWAS illuminates the molecular
415 basis of pleiotropy and genetic correlation. *Elife* 2022;11. DOI: 10.7554/eLife.79348.
- 416 37. Sun T, Chen Q, Li Y. Associations of serum cystatin C with depressive symptoms and suicidal ideation in major
417 depressive disorder. *BMC Psychiatry* 2021;21(1):576. DOI: 10.1186/s12888-021-03509-3.
- 418 38. Dastych M, Synek O, Gottwaldova J. Impact of Long-Term Lithium Treatment on Renal Function in Patients With Bipolar
419 Disorder Based on Novel Biomarkers. *J Clin Psychopharmacol* 2019;39(3):238-242. DOI: 10.1097/JCP.0000000000001030.
- 420 39. Campbell IH, Campbell H, Smith DJ. Insulin signaling as a therapeutic mechanism of lithium in bipolar disorder. *Transl*
421 *Psychiatry* 2022;12(1):350. DOI: 10.1038/s41398-022-02122-6.
- 422 40. Gabbouj S, Ryhänen S, Marttinen M, et al. Altered Insulin Signaling in Alzheimer's Disease Brain – Special Emphasis on
423 PI3K-Akt Pathway. *Frontiers in Neuroscience* 2019;13 (Mini Review) (In English). DOI: 10.3389/fnins.2019.00629.
- 424 41. Moorhead TWJ, McKirdy J, Sussmann JED, et al. Progressive Gray Matter Loss in Patients with Bipolar Disorder.
425 *Biological Psychiatry* 2007;62(8):894-900. DOI: <https://doi.org/10.1016/j.biopsych.2007.03.005>.
- 426 42. Huang J, Tao Q, Ang TFA, et al. The impact of increasing levels of blood C-reactive protein on the inflammatory loci SPI1
427 and CD33 in Alzheimer's disease. *Transl Psychiatry* 2022;12(1):523. DOI: 10.1038/s41398-022-02281-6.
- 428 43. Hanlon P, Jani BD, Nicholl B, Lewsey J, McAllister DA, Mair FS. Associations between multimorbidity and adverse health
429 outcomes in UK Biobank and the SAIL Databank: A comparison of longitudinal cohort studies. *PLoS Med*
430 2022;19(3):e1003931. DOI: 10.1371/journal.pmed.1003931.
- 431 44. Brayne C, Moffitt TE. The limitations of large-scale volunteer databases to address inequalities and global challenges in
432 health and aging. *Nat Aging* 2022;2(9):775-783. DOI: 10.1038/s43587-022-00277-x.
- 433 45. Glynn RJ. Promises and limitations of mendelian randomization for evaluation of biomarkers. *Clin Chem*
434 2010;56(3):388-90. DOI: 10.1373/clinchem.2009.142513.
- 435 46. Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey Smith G. Best (but oft-forgotten) practices: the design,
436 analysis, and interpretation of Mendelian randomization studies. *Am J Clin Nutr* 2016;103(4):965-78. DOI:
437 10.3945/ajcn.115.118216.
- 438 47. Smith GD, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. *Int J Epidemiol* 2004;33(1):30-42.
439 DOI: 10.1093/ije/dyh132.

440 **Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual
441 author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury
442 to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Exposure || Outcome



alpha

● 0.5

method

● Inverse variance weighted

● MR Egger

● Simple mode

● Weighted median

● Weighted mode