Mendelian Randomization Analysis Using Multiple Biomarkers of an Underlying Common Exposure

Jin Jin[1], Guanghao Qi[2], Zhi Yu[3,4,5], and Nilanjan Chatterjee[1,6,*]

[1]Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins Univeristy, Baltimore, Maryland, U.S.

[2]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland, U.S.

[3]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, U.S.

[4]Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts, U.S.

[5]Center for Genomic Medicine, Massachusetts General Hospital, Boston, Massachusetts, U.S.

[6]Department of Oncology, School of Medicine, Johns Hopkins University, Baltimore, Maryland, U.S.

[*]*nilanjan@jhu.edu*

Abstract

Mendelian Randomization (MR) analysis is increasingly popular for testing the causal effect of exposures on disease outcomes using data from genome-wide association studies. In some settings, the underlying exposure, such as systematic inflammation, may not be directly observable, but measurements can be available on multiple biomarkers, or other types of traits, that are co-regulated by the exposure. We propose method MRLE, which tests the significance for, and the direction of, the effect of a latent exposure by leveraging information from multiple related traits. The method is developed by constructing a set of estimating functions based on the second-order moments of summary association statistics, under a structural equation model where genetic variants are assumed to have indirect effects through the latent exposure and potentially direct effects on the traits. Simulation studies showed that MRLE has well-controlled type I error rates and increased power compared to single-trait MR tests under various types of pleiotropy. Applications of MRLE using genetic association statistics across five inflammatory biomarkers (CRP, IL-6, IL-8, TNF-α and MCP-1) provided evidence for potential causal effects of inflammation on increased risk of coronary artery disease, colorectal cancer and rheumatoid arthritis, while standard MR analysis for individual biomarkers often failed to detect consistent evidence for such effects.

Introduction

In the last two decades, Mendelian Randomization (MR) analysis has become increasingly popular for investigating the causal effects of risk factors and biomarkers on disease outcomes when randomized controlled trials are unavailable.[1–3] MR is a type of instrumental variable (IV) analysis which uses genetic variants (SNPs) as "instruments" aiming for unbiased estimate of the underlying causal effect when there exist potential confounders. While early MR analysis focused on using individual genetic variants of known functional consequence as instruments,[2,4–6] the era of genome-wide association studies (GWAS) has led to the rise of powerful MR analysis method based on mutiple SNPs.[7–10] In particular, with the rapid development of large publicly available GWAS, the majority of MR research has been transferred from working with individual-level data to two-sample MR analysis with summary-level data.[11–14]

Many recent methods for MR analysis have focused on improving robustness to the presence of pleiotropic association by which genetic instruments can affect the outcome of interest independent of the exposure and thus leading to violation of key assumptions.[15–23] In this study, we consider a novel setting of pleiotropic association, where genetic variants might be associated with multiple observed traits through an underlying latent exposure which may have a causal effect on the outcome. An example of the relevant issues has recently been illustrated in a study of the causal relationship between blood pressure and multiple kidney function biomarkers.[24] The analysis showed that standard MR analysis using individual kidney function biomarker, such as the estimated glomerular filtration rate determined based on the serum creatinine (eGFRcr), failed to detect a causal effect of kidney function on blood pressure. Instead, when the analysis was restricted to instruments which showed association across multiple biomarkers and thus were likely to be related to the underlying kidney function, clear evidence of causal effects emerged. Examples of groups of traits that may be governed by an underlying common exposure appear in many settings, including but not limited to, inflammatory biomarkers,[25,26] groups of metabolites related to dietary and lifestyle exposures,[27,28] and measurement instruments for evaluations of

underlying mental health conditions.[29,30]

We propose a novel method for conducting MR analysis for Latent Exposure (MRLE), which allows testing for the statistical significance and direction of the effect of an unobservable latent exposure by leveraging information from multiple related observable traits. The method uses summary-level association statistics of the observable traits and the outcome on a set of "strictly selected" SNPs (IVs) that are associated with at least two of the traits at specified thresholds of significance. We use an underlying structural equation model to describe causal paths between the SNPs, the latent exposure, the traits co-regulated by the exposure, and the outcome. or indirect association with the SNPs through its association with the latent exposure. We then construct a series of estimating functions by equating the second-order sample moments of the summary-level association statistics with the corresponding theoretical moments, and propose inference for identifiable parameters based on Generalized Method of Moments theory.[31–33]

We show by simulation that the proposed MRLE test has a well-controlled type I error rate, increased power as well as higher probability of correctly identifying the direction of the causal effect compared to single-trait MR analysis, and is robust to the presence of various types of pleiotropy across the traits and the outcome. We applied MRLE to test the effect of chronic inflammation on rheumatoid arthritis, coronary artery disease, and cancers including colorectal cancer, prostate cancer and endometrial cancer, using up to five inflammatory biomarkers including c-reactive protein (CRP), Interleukin 6 (IL-6), Interleukin 8 (IL-8), tumor necrosis factor alpha (TNF-α), and monocyte chemoattractant protein-1 (MCP-1). These analyses revealed that while MR analyses based on single biomarkers often failed to detect any consistent patterns of causal effects, the MRLE method detected such evidence for a number of diseases including RA, CAD and CRC.

Subjects and Methods

Problem setting and model setup

Let X denote an exposure variable of interest for which we want to examine its causal effect on an outcome Y, which we denote by θ. We assume that measurements on X are not directly available, and instead one can observe a set of biomarkers, or other types of traits, $\{B_k, k = 1, \ldots, K\}$, $K \geq 2$, that are co-regulated by X with known directions. We call a biomarker "valid" for such analysis on the exposure X if it meets the following two conditions: (A) X affects the biomarker, and (B) Conditional on X, the biomarker does not have an effect on the outcome itself ("pure surrogate"). Here we will show the development of the method for a continuous outcome Y, but the testing result is invariant with respect to the scale of the effect size, and thus the methods could be directly applied to other types of disease outcomes without modification (please see web appnedix for details).

Suppose we select M independent SNPs, G_j, $j = 1, 2, \ldots, M$, that are associated with one or more of the B_ks as instrumental variables (IVs) for the MR analysis. Details of the IV selection strategy will be discussed later in the section "Selection of genetic IVs". Following we will assume that the outcome Y is continuous and potentially transformed to be approximately normally distributed, and develop our method in the linear regression framework. We further assume without loss of generality, that the SNPs G_j, $j = 1, 2, \ldots, M$, the biomarkers B_k, $k = 1, 2, \ldots, K$, the latent exposure X, and the outcome Y are all standardized to have unit variance, so that the effect θ is on this standardized scale.

Figure 1 describes the assumed causal paths between the SNPs (G_js), the latent exposure of interest (X), the biomarkers co-regulated by the latent exposure (B_ks), and the outcome (Y), assuming a total of $K = 3$ biomarkers. Here we assume an "outcome" model of the form

$$Y = \theta X + \epsilon_y,$$

where ϵ_y denotes a mean-zero error term with variance $\sigma^2_{\epsilon_y}$, and given the unit variance of X and Y, we have $\theta^2 + \sigma^2_{\epsilon_y} = 1$. The model implies that given the true exposure X, all of the SNPs and biomarkers are independent of the outcome. Our goal is to conduct a hypothesis test with null hypothesis $H_0 : \theta = 0$ and alternative hypothesis $H_a : \theta \neq 0$, also to infer the sign of θ when the null hypothesis is rejected.

Next we assume an "exposure" model of the form

$$X = \sum_{j=1}^{M} \beta_{x,j} G_j + \epsilon_x,$$

where $\beta_{x,j}$, $j = 1, 2, \ldots, M$, denote additive effects of SNPs on X, and ϵ_x is a mean-zero error term with variance $\sigma^2_{\epsilon_x}$. We further assume that the effect sizes $\beta_{x,j}$, $j = 1, 2, \ldots, M$, can be modelled as mean-zero random variables with variance h^2_x. Given the unit variance of G_js and X, we have $Mh^2_x + \sigma^2_{\epsilon_x} = 1$.

Finally, we assume a set of "biomarker" models in the form

$$B_k = \theta_k X + \sum_{j=1}^{M} \gamma_{k,j} G_j + \epsilon_{B_k},$$

where θ_k denotes the effect of exposure X on biomarker B_k, $\gamma_{k,j}$, $j = 1, \ldots, M$, denote the direct effects of the SNPs on B_k that are not medicated through X, and ϵ_{B_k} denotes the mean-zero residual error term associated with B_k assumed to have a variance $\sigma^2_{\epsilon_{B_k}}$. We further assume that for each k, $\gamma_{k,j}$, $j = 1, \ldots, M$, are independent, mean-zero random variables with variance $h^2_{\gamma_k}$, and are also uncorrelated with association coefficients for the SNPs associated with X, i.e. $\beta_{x,j}$, $j = 1, \ldots, M$, in the "exposure" model.

We assume that we have summary-level association statistics for biomarker B_k, $\{(\widehat{\beta}_{B_k,j}, \sigma^2_{B_k,j})$, $j = 1, 2, \ldots, M\}$, and those for the outcome Y, $\{(\widehat{\beta}_{Y,j}, \sigma^2_{Y,j}), j = 1, 2, \ldots, M\}$, from separate GWAS, where $\widehat{\beta}$ and σ^2 denote the estimated association coefficient and its estimated standard

error, respectively. We allow the GWAS for B_ks to have potentially overlapping samples, but assume that the GWAS for Y is independent from those for the B_ks.[34]

It has been common practice to test for causal effects for an underlying exposure, such as chronic inflammation, using a surrogate biomarker such as CRP. Under the above structural equation model, we can derive asymptotic bias of standard IVW estimator in the form, $\theta\left(\theta_B h_x^2/\left(\theta_B^2 h_x^2 + h_\gamma^2\right) - 1\right)$. Thus, under the above model, while standard test based on individual biomarker is expected to be valid, it can lose major power due to attenuation of the underlying causal effect. Further, as we will show later, in the presence of more complex pleiotropy, single biomarker based test for causal effects can easily become invalid due to increased type-I error rates.

Cross-biomarker MR analysis using generalized method of moments

We propose statistical inference based on method of moments which only requires moments of the summary-level association statistics and no additional distributional assumptions for the model parameters. Let $X^s = (X_1^s, X_2^s, \ldots, X_M^s)^T$ denote the observed GWAS summary-level data, with $X_j^s = \left(\{\widehat{\beta}_{B_k,j}\}_{k=1}^K, \{\sigma_{B_k,j}^2\}_{k=1}^K, \widehat{\beta}_{Y,j}, \sigma_{Y,j}^2\right)^T$, $j = 1, 2, \ldots, M$. We consider the second-order moments of $\{\widehat{\beta}_{Y,j}, \widehat{\beta}_{B_1,j}, \ldots, \widehat{\beta}_{B_K,j}\}$, which, based on the model described in the previous section, can be derived as

$$
\begin{aligned}
\mathrm{E}(\widehat{\beta}_{Y,j}\widehat{\beta}_{B_k,j}) &= \theta\theta_k h_x^2, \quad 1 \le k \le K, \\
\mathrm{E}(\widehat{\beta}_{B_k,j}^2) &= \theta_k^2 h_x^2 + h_{\gamma_k}^2 + \sigma_{B_k,j}^2, \quad 1 \le k \le K, \\
\mathrm{E}(\widehat{\beta}_{B_k,j}\widehat{\beta}_{B_{k'},j}) &= \theta_k\theta_{k'} h_x^2 + c_{k,k'}, \quad 1 \le k < k' \le K,
\end{aligned}
$$

where $c_{k,k'}$ denotes the covariance between $\widehat{\beta}_{B_k,j}$ and $\widehat{\beta}_{B_{k'},j}$, which could arise due to sample overlap between GWAS of correlated biomarkers, and can be estimated from summary-statistics data using techniques such as bivariate LD score regression.[35,36] Detailed derivations are provided

in the Supplemental Data.

To avoid identifiability issue, we reparameterize the model with $\mu = \theta h_x$, $\mu_k = \theta_k h_x$, $k = 1, 2, \ldots, K$, and denote the vector of model parameters as $\eta = (\mu, \mu_1, \ldots, \mu_K, h^2_{\gamma_1}, \ldots, h^2_{\gamma_K})^T$. By equating the second-order moments of the summary-level association coefficients with the corresponding sample moments, we then define our estimating equations, $\Psi(X^s, \eta) = (1/M) \sum_{j=1}^{M} \psi(X^s_j, \eta) = 0$, where $\psi(X^s_j, \eta) = \Big(\psi_1(X^s_j, \eta), \psi_2(X^s_j, \eta), \ldots, \psi_d(X^s_j, \eta)\Big)^T$ with

$$\psi_k(X^s_j, \eta) = \widehat{\beta}_{Y,j}\widehat{\beta}_{B_k,j} - \mu_k\mu, \quad 1 \le k \le K,$$
$$\psi_{K+k}(X^s_j, \eta) = \widehat{\beta}^2_{B_k,j} - \sigma^2_{B_k,j} - h^2_{\gamma_k} - \mu^2_k, \quad 1 \le k \le K,$$
$$\psi_{\nu(k,k')}(X^s_j, \eta) = \widehat{\beta}_{B_k,j}\widehat{\beta}_{B_{k'},j} - c_{k,k'} - \mu_k\mu_{k'}, \quad 1 \le k < k' \le K,$$

where $\nu(k, k') = 2K + \sum_{l=1}^{k-1}(K - l) + (k' - k)$ and $d = K(K + 5)/2$. The inference on θ is then equivalent to that on μ, i.e. we test for $H^*_0 : \mu = 0$ versus $H^*_a : \mu \neq 0$, and determine the sign of θ from that of μ under the assumption that the directions of association of B_ks with X, i.e., sign of θ_ks, are known.

We observe that there are $d = K(K + 5)/2$ equations and $p = 2K + 1$ unknown parameters. To solve the estimating equations, we need $d \ge p$, i.e. a minimum of $K \ge 2$ biomarkers are required. When $K = 2$, the number of estimating equations equals the number of unknown parameters, and we can solve the equations directly. When $K \ge 2$, there are more equations than unknown parameters, the problem becomes over-identified and we may not be able to solve the exact equations. To handle this issue, we consider the generalized method of moments (GMM), and define the following GMM estimator,

$$\widehat{\eta}^{\text{GMM}} = \underset{\eta}{\operatorname{argmin}}\, Q_{\widehat{W}}(\eta),$$

which is the global minimizer of the objective function $Q_{\widehat{W}}(\eta) = \Psi(X^s, \eta)^T \widehat{W} \Psi(X^s, \eta)$, where

$\widehat{W}$ is a positive semi-definite weighting matrix which is typically specified based on the data X^s. Inference on the causal effect θ is then conducted based on $\widehat{\mu}^{\text{GMM}} = \widehat{\eta}^{\text{GMM}}_{(1)}$. We obtain $\widehat{\mu}^{\text{GMM}}$ and an estimate of its standard error based on GMM theory, and then perform Wald test on $H_0 : \mu = 0$ versus $H_a : \mu \neq 0$. Properties of $\widehat{\mu}^{\text{GMM}}$ and computational details are summarized in the Supplemental Data.

Selection of genetic IVs

A key issue of simultaneously analyzing multiple biomarkers is how to select instruments that may be associated with the underlying common exposure. As noted earlier, many SNPs that are associated with individual biomarkers may represent genetic variations that are unrelated to the exposure and use of these may lead to loss of power. Further, if there are numerous biomarkers, the chance of exposure-independent pleiotropic association of SNPs with some of the biomarkers also increases, potentially leading to bias.

To alleviate these issues, we propose a more strict IV selection procedure, where we select SNPs that are associated with at least two of the biomarkers co-regulated by the latent exposure. Intuitively, SNPs which are associated with multiple biomarkers are more likely to be related to the underlying common exposure and is less likely to suffer from pleiotropic effects due to any shared genetic background between individual biomarkers and the outcome. However, the use of GWAS significance threshold (5×10^{-8}) can be overly strict for this IV selection strategy especially when GWAS for some biomarkers have relatively small sample sizes. For example, in our data analysis, which will be described later in detail, while GWAS for CRP was large ($N = 320041$), but those for the other biomarkers were relatively small ($N = 3454 \sim 8394$). To ensure that a minimum number of IVs are selected, a more liberal threshold can be used depending on the GWAS sample size.

Simulations assuming no pleiotropy

We illustrate performance of the proposed MRLE method by simulation studies under a variety of data scenarios. Assume that there are $M = 2 \times 10^5$ independent common SNPs across the whole genome, and that the summary-level data are available for $K = 4$, 6 or 8 biomarkers representing an underlying latent exposure. Instead of assuming all SNPs to be causal as in the "exposure" and "biomarker" model, we considered a more realistic setting where only a small proportion of the SNPs are causal, i.e. have non-zero effect on the traits. Specifically, we assumed that a random subset of $M_x = \pi_x M$ SNPs are associated with X, with effect sizes $\beta_{x,j}$s generated from $N(0, h_x^2)$; similarly, a random subset of $M_{B_k} = \pi_{B_k} M$ SNPs are associated with B_k, with effect sizes $\gamma_{k,j}$s generated from $N(0, h_{B_k}^2)$, $k = 1, 2, \ldots, K$. We set the proportion of causal SNPs to $\pi_x = \pi_{B_k} = 1\%$, and θ_k to $\sqrt{0.3}$, $k = 1, 2, \ldots, K$, , so that X explains 30% of variability of each of the biomarkers. We further set the total genetic heritability of each biomarker ($H^2_{\text{B.total}}$) to 0.2 or 0.3, and the proportion explained by the association with X to 0.2 or 0.3, which leads to a total heritability of each biomarker that is explained by X ($H^2_{B.X}$) between 0.04 and 0.09.

We directly simulated GWAS summary-level association statistics instead of individual-level data to reduce computation time (please see the Supplemental Data for details).[36] For simplicity, we assumed equal sample sizes across all GWAS and all SNPs, which was set equal to 6×10^4, 8×10^4 or 10^5. We also set the overlapping GWAS sample size between any two biomarkers to $N_{B_k,B_l} = N$, i.e. the GWAS summary-level data for all Biomarkers were obtained from the same set of individuals, and between-biomarker correlation to $cov(B_k, B_l) = 0.3$, $k, l \in 1, \ldots, K$. We also set $N_{B_k,Y} = 0$, $k = 1, \ldots, K$ based on the requirement of no sample overlap between the biomarkers and the outcome.

After generating summary-level association statistics, we selected IVs using either one of the two strategies discussed in Section , i.e. we selected (1) the union of the SNPs that reached genome-wide significance threshold (5×10^{-8}) for each B_k, the corresponding test was denoted as "MRLE-Union"; or (2) SNPs that reached a more liberal significance threshold, 5×10^{-4},

for at least two biomarkers, the corresponding test was denoted as "MRLE-Intersection". As a comparison, we also applied the fixed-effect IVW tests based on SNPs associated with each individual biomarker. Similiar to MRLE, we also considered two IV selection strategies, where for each single-biomarker IVW test, we selected (1) SNPs that reached genome-wide significance threshold (5×10^{-8}) for that biomarker only (IVW-Standard); or (2) SNPs that reached significance level 5×10^{-4} for that biomarker and at least one other (IVW-Intersection). Since multiple IVW tests were conducted, an adjusted significance level, $\alpha_1 = 1 - (1 - \alpha_0)^{1/K}$ was used for the test on each biomarker to control the family-wise error rate at α_0.

Simulations assuming pleiotropy between biomarkers and the outcome

We also considered various types of pleiotropic effects across biomarkers and the outcome independent of the exposure of interest X. We first considered pleiotropy between the biomarkers and the outcome, i.e. SNPs having correlated associations with the outcome and at least one of the biomarkers (Figure 2, panel A). In our simulation, this was reflected by randomly assigning half of the $\pi_B M$ SNPs that had a direct effect on each biomarker to have another direct effect, $u_{k,j}$, on the outcome, with mean 0, variance $0.1h^2_{B_k}$, and $\text{cor}(\gamma_{B_k,j}, u_{k,j}) = 0.15$, $k \in 1, \ldots, K$. We set $\pi_B = 2\%$, and for each biomarker, we set the total genetic heritability to 0.3 or 0.4, and the proportion of genetic heritability explained by the latent exposure to 0.15 or 0.2, which leads to a heritability of the biomarker explained by the latent exposure between 0.045 and 0.08.

Simulations assuming pleiotropy across biomarkers

Another type of pleiotropy that is likely to exist is the pleiotropy among biomarkers, i.e. SNPs having correlated direct effects across multiple biomarkers, which can be due to shared genetic pathways across biomarkers (Figure 2, panel B). In the simulation, we introduced this type of pleiotropy by allowing $1/K$ of the $\pi_B M$ SNPs that had a direct effect on biomarker k ($\gamma_{B_k,j}$) to also have another direct effect on each of the other $K - 1$ biomarkers ($\gamma_{B_l,j}$, $j \neq k$), with mean

0, variances h_B^2, and $\text{cor}(\gamma_{B_k,j}, \gamma_{B_l,j}) = 0.5$, $k, l \in 1, \ldots, K$. This leads to a total of $\pi_B M/K$ SNPs to be directly associated with each individual biomarker, and $\pi_B M/K$ SNPs to be directly associated with each pair of biomarkers. We set $\pi_B = K\%$, and for each biomarker, we set the total genetic heritability to 0.3 or 0.4, and the proportion of genetic heritability explained by the latent exposure to 0.15 or 0.2, which leads to a heritability of the biomarker explained by the latent exposure between 0.045 and 0.08.

Chronic inflammation on various diseases

Chronic inflammation has been long hypothesized to be one of the underlying causes for a spectrum of common diseases.[37] Epidemiologic studies have used a variety of biomarkers to study potential relationship between inflammation and diseases.[25,38–41] In particular, C-reactive protein (CRP), a type of protein in blood produced by liver in response to inflammation, has been commonly used to associate chronic inflammation to a variety of diseases including heart disease,[42,43] ischemic stroke,[44,45] cancers of colon[46,47] and lung.[48,49] However, recent MR studies have indicated that CRP itself is unlikely to be an underlying causal risk factor for these diseases. In addition to CRP, a variety of other biomarkers, including Interleukin 6 (IL-6), Interleukin 8 (IL-8), tumor necrosis factor alpha (TNF-α), and monocyte chemoattractant protein-1 (MCP-1), are commonly used to assess inflammation and hence associate with risks of diseases. MR analyses for these additional biomarkers, however, have been largely limited as sample sizes for the underlying GWAS have been typically fairly modest.[50–53]

Here we investigate the causal effect of chronic inflammation on a number of diseases including rheumatoid arthritis (RA), coronary artery disease (CAD), colorectal cancer (CRC), prostate cancer (PCa) and endometiral cancers (EC), all of which have been associated with one or more inflammatory biomarkers in previous studies.[43,54–64]

We apply the proposed MRLE method to test the causal effect of chronic inflammation on the diseases using summary-level data from publicly available GWAS for five commonly used

systematic biomarkers of chronic inflammation, including, CRP, IL-6, IL-8, TNF-α and MCP-1. We ourselves generated the summary-level data for CRP by conducting a GWAS on 320041 unrelated, European-ancestry individuals in the UK Biobank who have CRP measurements available. We performed a GWAS across 1186957 common SNPs (minor allele frequency >5%) that are available in HapMap 3[65] based on additive genetic model adjusting for age, sex and body-mass index (BMI) using PLINK 2.[66,67] We used summary-level data for IL-6, IL-8, TNF-α and MCP-1 which were previously generated based on GWAS on up to 3596 European-ancestry participants in the Cardiovascular Risk in Young Finns Study (YFS) and up to 6313 European-ancestry participants in the FINRISK study that have the corresponding measurements available, after adjusting for age, sex, BMI and the first 10 genetic principal components.[50] Summary-level data for RA,[68] CAD,[69] CRC,[70] PCa,[71] and EC[72] were all obtained from publicly available GWAS. As GWAS for CRC and EC have overlapping individuals with the UK Biobank-based GWAS for CRP, we excluded CRP from the set of biomarkers used in the analyses of CRC and EC. Detailed information on GWAS for the inflammatory biomarkers and diseases are summarized in Table S1.

To select IVs, we first conducted a filtering procedure[73] by removing the SNPs that were strand-ambiguous, had alleles that did not match those in the 1000 Genomes Project, or were within the major histocompatibility complex (MHC) region (26Mb - 34Mb on chromosome 6) since they may have complex large pleiotropic effects across multiple inflammation related traits.[74,75] We then selected SNPs that were significantly associated with at least two of the inflammatory biomarkers. Considering the GWAS sample sizes, we used a more liberal instrument selection threshold, $\alpha = 10^{-3}$, for the four cytokine-type biomarkers ($N_{\text{GWAS}} = 3454 \sim 8293$), and a more stringent threshold, $\alpha = 5 \times 10^{-6}$, for CRP ($N_{\text{GWAS}} = 320041$). To select independent SNPs, we conducted linkage disequilibrium (LD) clumping on the remaining SNPs with a window size $d = 1$MB and a cut-off for squared-correlation, $r^2 = 0.05$ using PLINK.[76] Additionally, we removed SNPs that were significantly associated ($\alpha = 5 \times 10^{-8}$) with the potential confounders

of the association between inflammation and the outcomes,[77] including systolic blood pressure (SBP), history of diabetes, smoking status, alcohol consumption status, high-density lipoprotein (HDL), and low-density lipoprotein (LDL), based on a GWAS we conducted on 320041 relatively unrelated European-ancestry UK Biobank individuals using PLINK 2.[66,67] These steps lead to a total of 53-58 IVs selected for the MR analyses on CRC and EC, and a total of 58-67 IVs for the MR analyses on CAD, RA and PCa.

Since summary-level data for the four cytokine-type biomarkers, IL-6, IL-8, TNF-α and MCP-1, were obtained from the same GWAS, we estimated the between-biomarker covariance ($c_{k,k'}$s in the proposed estimating equations) by fitting bivariate LD score regressions as described in Subject and Methods section. We also conducted fixed-effect IVW test on each biomarker, where the IVs were defined as the SNPs that were associated with the biomarker and at least one other biomarker. The MR analyses were conducted at a significance level of $\alpha_0 = 0.05$. The desired level of type I error rate for the single-biomarker IVW tests were set at $1-(1-\alpha_0)^{1/K}$, i.e. 0.0102 for CAD, RA and PCa, and 0.0127 for CRC and EC, to control the family-wise error rate of the IVW test at 0.05 for each disease. Details of the IV selection procedure are summarized in Figure S5.

Results

Simulation results assuming no pleiotropy

Under the no pleiotropy assumption, type I error rates seem to be well controlled at approximately $\alpha_0 = 0.05$ in all simulated settings (Figures 3, S1 and S2). We observe that overall, as the GWAS sample size (N) increases (which leads to increased number of IVs), both IVW and MRLE tests have increased power (Figure 3). The power of the tests also increases as the total genetic heritability of the biomarkers ($H^2_{\text{B.total}}$) and the proportion of this heritability explained by the exposure ($H^2_{\text{B.X}}/H^2_{\text{B.total}}$) increase. For both IVW and MRLE tests, the probability of correctly

identifying the direction of the effect becomes approximately 100% if the power exceeds 0.25 (Figure S3). Compared to the naive IV selection strategy of choosing all SNPs associated with any biomarker, the more strict strategy of choosing SNPs associated with at least two biomarkers yields a much high power and a higher probability of correctly identifying the direction of the effect for both tests (Figures 3, S1 and S2). Using either of the two IV selection strategies, MRLE provides a substantially higher power and a higher probability of correctly identifying the causal direction compared to the IVW test under the same level of type I error control. We also conducted simulation studies under the same settings assuming a total of $K = 4$ or 8 biomarkers (see Figures S1 and S2). Overall, the results are similar to those presented in Figures 3, but using a larger number of biomarkers tends to give a higher power and a higher chance of correctly identifying the causal direction.

Simulation results assuming pleiotropy between biomarkers and the outcome

In the presence of horizontal pleiotropy between a biomarker and the outcome, selecting SNPs significantly associated ($\alpha = 5 \times 10^{-8}$) with any biomarker as IVs can lead to severe inflation in type I error rate, especially for IVW test (Figure 4, panel A). Compared to the results in no-pleiotropy scenario in Figure 3, at the same level of power, the probability of correctly identifying the causal direction also decreases for both methods (Figure S4), although MRLE still outperforms IVW test in terms of type I error control and identification of the direction of the effect. On the contrary, taking SNPs that are associated with at least two biomarkers at a more liberal threshold ($\alpha = 5 \times 10^{-4}$) as IVs can substantially improve the type I error control, power, and identification of the direction of the effect for both methods. Under this strategy, MRLE has a consistently lower type I error rate that is close to the target level at $\alpha_0 = 0.05$ and higher power compared to IVW test.

Simulation results assuming pleiotropy across biomarkers

Simulation results show that in the presence of pleiotropy across biomarkers, strict type I error control can be achieved by either selecting SNPs associated with any biomarker ("IVW-Standard" and "MRLE-Union" in Figure 5, panel A) or selecting SNPs associated with at least two biomarkers ("IVW-Intersection" and "MRLE-Intersection" in Figure 5, panel A). Overall, the second IV selection strategy yields a higher power and a higher probability of correctly identifying the causal direction for both IVW and MRLE (Figures 5 and S4). Compared to IVW, MRLE provides a higher power and a higher probability of correctly identifying the causal direction using either of the two IV selection approaches.

The role of chronic inflammation in disease progression

We observe that results on RA are significant based on all tests (Table). The proposed MRLE test detects a significant, positive effect of chronic inflammation (p-value=$5.1 \times 10-27$). All single-biomarker IVW tests detect significant evidence as well, but the identified causal directions do not agree with each other: the tests based on IL-8, TNF-α and CRP suggest a positive effect of chronic inflammation on the risk of RA, while IL-6 and MCP-1 suggest a negative effect, thus no universal conclusion can be drawn for the effect on RA based on IVW tests. No single-biomarker IVW test shows significant evidence for a causal effect of chronic inflammation on the risk of CAD. Among the single-biomarker IVW tests, three of the biomarkers (CRP, TNF-α and IL-8) which indicate most significant evidence all seem to be associated with an increased risk of CAD. Similarly, no single-biomarker IVW test on CRC suggests a significant effect of chronic inflammation, although two of them (IL-8 and TNF-α) achieve borderline significance both indicating an increased risk of CRC being associated with higher level of inflammation. The proposed MRLE method, on the other hand, indicates a significant, positive effect of chronic inflammation on the risk of both CAD (p-value=0.012) and CRC (p-value=0.011). Neither the IVW tests nor the MRLE test detects any significant effect of chronic inflammation on PCa or

EC.

Discussion

We propose a novel method for MR analysis for testing the causal effect of an unobservable latent exposure utilizing multiple traits co-regulated by the exposure. Through a set of extensive simulation studies and data analyses, we demonstrate that the proposed method overcomes various challenges associated with the standard MR analyses that use individual observable traits associated with the latent exposure and their associated genetic instruments. The proposed MRLE test is applicable to either continous or dichotomous/categorical outcomes.

Several practical issues merit considerations. First, the validity of the selected observable traits as surrogates for the latent exposure of interest. Theoretically, including more observable traits for the latent exposure can provide higher power. However, the inclusion of invalid traits, i.e. traits that are actually not regulated by the latent exposure or/and themselves have direct causal effect on the exposure, can affect both type-I error rate and power of the tests. Second, a strict IV selection procedure is crucial. Other than the commonly implemented filtering procedures such as removing SNPs that are significantly associated with the potential confounders, we recommend selecting the SNPs that are associated with at least two traits. We have shown by simulation that this more strict criterion can efficiently reduce the number of invalid IVs selected, thus providing a higher power and stricter type I error control under various types of pleiotropy. Additionally, although a more liberal significance threshold may be used for an individual trait when the sample size is relatively small, there needs to be rigorous criteria to ensure selection of valid instruments. Recent methods for the discovery of strictly pleiotropic association across multiple traits can be potentially used for instrument selection.[78]

Results from our MR analysis of chronic inflammation and various diseases should be interpreted cautiously. First, null results for certain diseases, such as PCa and EC, may be due to the inability of the set of the biomarkers used to capture relevant aspects of inflammation. Second,

sample size of the GWAS for all the inflammatory biomarkers except CRP was relatively small, causing substantially large uncertainty. In the future, our results need to be confirmed using results from much larger GWAS of inflammatory biomarkers when such data become available.

Supplemental Data

Supplemental Data include detailed derivations of the formulas and equations, 5 Supplemental figures and 1 Supplemental table.

Acknowledgements

The UK Biobank data was accessed via application ID 17712. Research of Drs. Jin Jin, Guanghao Qi and Nilanjan Chatterjee was supported by an R01 grant from the National Human Genome Research Institute [1 R01 HG010480-01].

Declaration of Interests

The authors declare no competing interests.

References

[1] George Davey Smith and Shah Ebrahim. 'mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *International journal of epidemiology*, 32(1):1–22, 2003.

[2] Debbie A Lawlor, Roger M Harbord, Jonathan AC Sterne, Nic Timpson, and George Davey Smith. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Statistics in medicine*, 27(8):1133–1163, 2008.

[3] Connor A Emdin, Amit V Khera, and Sekar Kathiresan. Mendelian randomization. *Jama*, 318(19):1925–1926, 2017.

[4] Sander Greenland. An introduction to instrumental variables for epidemiologists. *International journal of epidemiology*, 29(4):722–729, 2000.

[5] Marianne Abifadel, Mathilde Varret, Jean-Pierre Rabès, Delphine Allard, Khadija Ouguerram, Martine Devillers, Corinne Cruaud, Suzanne Benjannet, Louise Wickham, Danièle Erlich, et al. Mutations in pcsk9 cause autosomal dominant hypercholesterolemia. *Nature genetics*, 34(2):154–156, 2003.

[6] Jonathan C Cohen, Eric Boerwinkle, Thomas H Mosley Jr, and Helen H Hobbs. Sequence variations in pcsk9, low ldl, and protection against coronary heart disease. *New England Journal of Medicine*, 354(12):1264–1272, 2006.

[7] Brandon L Pierce, Habibul Ahsan, and Tyler J VanderWeele. Power and instrument strength requirements for mendelian randomization studies using multiple genetic variants. *International journal of epidemiology*, 40(3):740–752, 2011.

[8] Brandon L Pierce and Stephen Burgess. Efficient design for mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *American journal of epidemiology*, 178(7):1177–1184, 2013.

[9] David M Evans and George Davey Smith. Mendelian randomization: new applications in the coming age of hypothesis-free causality. *Annual review of genomics and human genetics*, 16:327–350, 2015.

[10] Jie Zheng, Denis Baird, Maria-Carolina Borges, Jack Bowden, Gibran Hemani, Philip Haycock, David M Evans, and George Davey Smith. Recent developments in mendelian randomization studies. *Current epidemiology reports*, 4(4):330–345, 2017.

[11] John R Thompson, Cosetta Minelli, Keith R Abrams, Martin D Tobin, and Richard D Riley. Meta-analysis of genetic studies using mendelian randomization—a multivariate approach. *Statistics in medicine*, 24(14):2241–2254, 2005.

[12] Leonelo E Bautista, Liam Smeeth, Aroon D Hingorani, and Juan P Casas. Estimation of bias in nongenetic observational studies using "mendelian triangulation". *Annals of epidemiology*, 16(9):675–680, 2006.

[13] Stephen Burgess and Simon G Thompson. Multivariable mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *American journal of epidemiology*, 181(4):251–260, 2015.

[14] Guanghao Qi and Nilanjan Chatterjee. A comprehensive evaluation of methods for Mendelian randomization using realistic simulations and an analysis of 38 biomarkers for risk of type 2 diabetes. *International Journal of Epidemiology*, 01 2021. dyaa262.

[15] Jack Bowden, George Davey Smith, and Stephen Burgess. Mendelian randomization with invalid instruments: effect estimation and bias detection through egger regression. *International journal of epidemiology*, 44(2):512–525, 2015.

[16] Jack Bowden, George Davey Smith, Philip C Haycock, and Stephen Burgess. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genetic epidemiology*, 40(4):304–314, 2016.

[17] Fernando Pires Hartwig, George Davey Smith, and Jack Bowden. Robust inference in summary data mendelian randomization via the zero modal pleiotropy assumption. *International journal of epidemiology*, 46(6):1985–1998, 2017.

[18] Marie Verbanck, Chia-yen Chen, Benjamin Neale, and Ron Do. Detection of widespread horizontal pleiotropy in causal relationships inferred from mendelian randomization between complex traits and diseases. *Nature genetics*, 50(5):693–698, 2018.

[19] Zhihong Zhu, Zhili Zheng, Futao Zhang, Yang Wu, Maciej Trzaskowski, Robert Maier, Matthew R Robinson, John J McGrath, Peter M Visscher, Naomi R Wray, et al. Causal associations between risk factors and common diseases inferred from gwas summary data. *Nature communications*, 9(1):1–12, 2018.

[20] Guanghao Qi and Nilanjan Chatterjee. Mendelian randomization analysis using mixture models for robust and efficient estimation of causal effects. *Nature communications*, 10(1):1–10, 2019.

[21] Jean Morrison, Nicholas Knoblauch, Joseph H Marcus, Matthew Stephens, and Xin He. Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. *Nature Genetics*, pages 1–7, 2020.

[22] Chin Yang Shapland, Qingyuan Zhao, and Jack Bowden. Profile-likelihood bayesian model averaging for two-sample summary data mendelian randomization in the presence of horizontal pleiotropy. *BioRxiv*, 2020. doi: https://doi.org/10.1101/2020.02.11.943712.

[23] Qingyuan Zhao, Jingshu Wang, Gibran Hemani, Jack Bowden, Dylan S Small, et al. Statistical inference in two-sample summary-data mendelian randomization using robust adjusted profile score. *Annals of Statistics*, 48(3):1742–1769, 2020.

[24] Zhi Yu, Josef Coresh, Guanghao Qi, Morgan Grams, Eric Boerwinkle, Harold Snieder, Alexander Teumer, Cristian Pattaro, Anna Köttgen, Nilanjan Chatterjee, et al. A bidirectional mendelian randomization study supports causal effects of kidney function on blood pressure. *Kidney International*, 2020.

[25] Darren R Brenner, Dominique Scherer, Kenneth Muir, Joellen Schildkraut, Paolo Boffetta, Margaret R Spitz, Loic Le Marchand, Andrew T Chan, Ellen L Goode, Cornelia M Ulrich, et al. A review of the application of inflammatory biomarkers in epidemiologic cancer research. *Cancer Epidemiology and Prevention Biomarkers*, 23(9):1729–1751, 2014.

[26] Shehua Qian, Olga Golubnitschaja, and Xianquan Zhan. Chronic inflammation: Key player and biomarker-set to predict and prevent cancer development and progression based on individualized patient profiles. *Epma Journal*, pages 1–17, 2019.

[27] Qianqian Gu, John J Spinelli, Trevor BJ Dummer, Treena E McDonald, Steven C Moore, and Rachel A Murphy. Metabolic profiling of adherence to diet, physical activity and body size recommendations for cancer prevention. *Scientific reports*, 8(1):1–11, 2018.

[28] Kolade Oluwagbemigun, Jana Foerster, Claire Watkins, Fiona Fouhy, Catherine Stanton, Manuela M Bergmann, Heiner Boeing, and Ute Nöthlings. Dietary patterns are associated with serum metabolite patterns and their association is influenced by gut bacteria among older german adults. *The Journal of nutrition*, 150(1):149–158, 2020.

[29] Athanasios Tsanas, Kate Saunders, Amy Bilderbeck, Niclas Palmius, Guy Goodwin, and Maarten De Vos. Clinical insight into latent variables of psychiatric questionnaires for mood symptom self-assessment. *JMIR mental health*, 4(2):e15, 2017.

[30] Louise Black, Margarita Panayiotou, and Neil Humphrey. The dimensionality and latent structure of mental health difficulties and wellbeing in early adolescence. *PloS one*, 14(2):e0213018, 2019.

[31] Lars Peter Hansen. Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054, 1982.

[32] KW Newey and Daniel McFadden. Large sample estimation and hypothesis. *Handbook of Econometrics, IV, Edited by RF Engle and DL McFadden*, pages 2112–2245, 1994.

[33] Alastair R Hall et al. *Generalized method of moments*. Oxford university press, 2005.

[34] Stephen Burgess, Neil M Davies, and Simon G Thompson. Bias due to participant overlap in two-sample mendelian randomization. *Genetic epidemiology*, 40(7):597–608, 2016.

[35] Patrick Turley, Raymond K Walters, Omeed Maghzian, Aysu Okbay, James J Lee, Mark Alan Fontana, Tuan Anh Nguyen-Viet, Robbee Wedow, Meghan Zacher, Nicholas A Furlotte, et al. Multi-trait analysis of genome-wide association summary statistics using mtag. *Nature genetics*, 50(2):229–237, 2018.

[36] Guanghao Qi and Nilanjan Chatterjee. Heritability informed power optimization (hipo) leads to enhanced detection of genetic associations across multiple traits. *PLoS genetics*, 14(10):e1007549, 2018.

[37] Peter Libby. Inflammatory mechanisms: the molecular basis of inflammation and disease. *Nutrition reviews*, 65(suppl_3):S140–S146, 2007.

[38] Philip Hunter. The inflammation theory of disease: The growing realization that chronic inflammation is crucial in many diseases opens new avenues for treatment. *EMBO reports*, 13(11):968–970, 2012.

[39] Jeanette M Bennett, Glenn Reeves, George E Billman, and Joachim P Sturmberg. Inflammation–nature's way to efficiently respond to all types of challenges: implications for understanding and managing "the epidemic" of chronic diseases. *Frontiers in Medicine*, 5:316, 2018.

[40] David Furman, Judith Campisi, Eric Verdin, Pedro Carrera-Bastos, Sasha Targ, Claudio Franceschi, Luigi Ferrucci, Derek W Gilroy, Alessio Fasano, Gary W Miller, et al. Chronic inflammation in the etiology of disease across the life span. *Nature medicine*, 25(12):1822–1832, 2019.

[41] Sedat Demir. The process of acute and chronic inflammation: Biomarkers and their relationship with diseases. In *Role of Nutrition in Providing Pro-/Anti-Inflammatory Balance: Emerging Research and Opportunities*, pages 1–23. IGI Global, 2020.

[42] Emerging Risk Factors Collaboration et al. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. *The Lancet*, 375(9709):132–140, 2010.

[43] Amit Kumar Shrivastava, Harsh Vardhan Singh, Arun Raizada, and Sanjeev Kumar Singh. C-reactive protein, inflammation and coronary heart disease. *The Egyptian Heart Journal*, 67(2):89–97, 2015.

[44] Mario Di Napoli, Francesca Papa, and Vittorio Bocola. C-reactive protein in ischemic stroke: an independent prognostic factor. *Stroke*, 32(4):917–924, 2001.

[45] Reyna L VanGilder, Danielle M Davidov, Kyle R Stinehart, Jason D Huber, Ryan C Turner, Karen S Wilson, Eric Haney, Stephen M Davis, Paul D Chantler, Laurie Theeke, et al. C-reactive protein and long-term ischemic stroke prognosis. *Journal of Clinical Neuroscience*, 21(4):547–553, 2014.

[46] Thomas P Erlinger, Elizabeth A Platz, Nader Rifai, and Kathy J Helzlsouer. C-reactive protein and the risk of incident colorectal cancer. *Jama*, 291(5):585–590, 2004.

[47] Krasimira Aleksandrova, Mazda Jenab, Heiner Boeing, Eugene Jansen, H Bas Bueno-de Mesquita, Sabina Rinaldi, Elio Riboli, Kim Overvad, Christina C Dahm, Anja Olsen, et al. Circulating c-reactive protein concentrations and risks of colon and rectal cancer: a nested case-control study within the european prospective investigation into cancer and nutrition. *American journal of epidemiology*, 172(4):407–418, 2010.

[48] Anil K Chaturvedi, Neil E Caporaso, Hormuzd A Katki, Hui-Lee Wong, Nilanjan Chatterjee, Sharon R Pine, Stephen J Chanock, James J Goedert, and Eric A Engels. C-reactive protein and risk of lung cancer. *Journal of clinical oncology*, 28(16):2719, 2010.

[49] Ugo Pastorino, Daniele Morelli, Giovanni Leuzzi, Mara Gisabella, Paola Suatoni, Francesca Taverna, Elena Bertocchi, Mattia Boeri, Gabriella Sozzi, Anna Cantarutti, et al. Baseline and

postoperative c-reactive protein levels predict mortality in operable lung cancer. *European journal of cancer*, 79:90–97, 2017.

[50] Ari V Ahola-Olli, Peter Würtz, Aki S Havulinna, Kristiina Aalto, Niina Pitkänen, Terho Lehtimäki, Mika Kähönen, Leo-Pekka Lyytikäinen, Emma Raitoharju, Ilkka Seppälä, et al. Genome-wide association study identifies 27 loci influencing concentrations of circulating cytokines and growth factors. *The American Journal of Human Genetics*, 100(1):40–50, 2017.

[51] Julia Höglund, Nima Rafati, Mathias Rask-Andersen, Stefan Enroth, Torgny Karlsson, Weronica E Ek, and Åsa Johansson. Improved power and precision with whole genome sequencing data in genome-wide association studies of inflammatory biomarkers. *Scientific reports*, 9, 2019.

[52] Robert F Hillary, Daniel Trejo-Banos, Athanasios Kousathanas, Daniel L McCartney, Sarah E Harris, Anna J Stevenson, Marion Patxot, Sven Erik Ojavee, Qian Zhang, David C Liewald, et al. Multi-method genome-and epigenome-wide studies of inflammatory protein levels in healthy older adults. *Genome medicine*, 12(1):1–15, 2020.

[53] Abigail Emma Russell, Tamsin Ford, David Gunnell, Jon Heron, Carol Joinson, Paul Moran, Caroline Relton, Matthew Suderman, Gibran Hemani, and Becky Mars. Investigating evidence for a causal association between inflammation and self-harm: a multivariable mendelian randomisation study. *Brain, behavior, and immunity*, 2020.

[54] Ernest HS Choy and Gabriel S Panayi. Cytokine pathways and joint inflammation in rheumatoid arthritis. *New England Journal of Medicine*, 344(12):907–916, 2001.

[55] Sarah Kraus and Nadir Arber. Inflammation and colorectal cancer. *Current opinion in pharmacology*, 9(4):405–410, 2009.

[56] Christine M Friedenreich, Annie R Langley, Thomas P Speidel, David CW Lau, Kerry S Courneya, Ilona Csizmadi, Anthony M Magliocco, Yutaka Yasui, and Linda S Cook. Case–control study of inflammatory markers and the risk of endometrial cancer. *European journal of cancer prevention*, 22(4):374–379, 2013.

[57] Monther Abu-Remaileh, Sebastian Bender, Günter Raddatz, Ihab Ansari, Daphne Cohen, Julian Gutekunst, Tanja Musch, Heinz Linhart, Achim Breiling, Eli Pikarsky, et al. Chronic inflammation induces a novel epigenetic program that is conserved in intestinal adenomas and in colorectal cancer. *Cancer research*, 75(10):2120–2130, 2015.

[58] Monika Izano, Esther K Wei, Caroline Tai, Helen Swede, Steven Gregorich, Tamara B Harris, Heidi Klepin, Suzanne Satterfield, Rachel Murphy, Anne B Newman, et al. Chronic inflammation and risk of colorectal and other obesity-related cancers: the health, aging and body composition study. *International journal of cancer*, 138(5):1118–1128, 2016.

[59] Elizabeth A Platz, Ibrahim Kulac, John R Barber, Charles G Drake, Corinne E Joshu, William G Nelson, M Scott Lucia, Eric A Klein, Scott M Lippman, Howard L Parnes, et al. A prospective study of chronic inflammation in benign prostate tissue and risk of prostate cancer: linked pcpt and select cohorts. *Cancer Epidemiology and Prevention Biomarkers*, 26(10):1549–1557, 2017.

[60] Hongyu Li, Kai Sun, Ruiping Zhao, Jiang Hu, Zhiru Hao, Fei Wang, Yaojun Lu, Fu Liu, and Yong Zhang. Inflammatory biomarkers of coronary heart disease. *Front Biosci (Schol Ed)*, 10:185–96, 2018.

[61] Paul M Ridker, Jean G MacFadyen, Brendan M Everett, Peter Libby, Tom Thuren, Robert J Glynn, John Kastelein, Wolfgang Koenig, Jacques Genest, Alberto Lorenzatti, et al. Relationship of c-reactive protein reduction to cardiovascular event reduction following treatment

with canakinumab: a secondary analysis from the cantos randomised controlled trial. *The Lancet*, 391(10118):319–328, 2018.

[62] Tommaso Cai, Raffaella Santi, Irene Tamanini, Ilaria Camilla Galli, Gianpaolo Perletti, Truls E Bjerklund Johansen, and Gabriella Nesi. Current knowledge of the potential links between inflammation and prostate cancer. *International journal of molecular sciences*, 20(15):3833, 2019.

[63] Isaac Subirana, Montserrat Fitó, Oscar Diaz, Joan Vila, Albert Francés, Eva Delpon, Juan Sanchis, Roberto Elosua, Daniel Muñoz-Aguayo, Irene R Dégano, et al. Prediction of coronary disease incidence by biomarkers of inflammation, oxidation, and metabolism. *Scientific reports*, 8(1):1–7, 2018.

[64] Xiaoliang Wang, James Y Dai, Demetrius Albanes, Volker Arndt, Sonja I Berndt, Stéphane Bézieau, Hermann Brenner, Daniel D Buchanan, Katja Butterbach, Bette Caan, et al. Mendelian randomization analysis of c-reactive protein on colorectal cancer risk. *International journal of epidemiology*, 48(3):767–780, 2019.

[65] International HapMap 3 Consortium et al. Integrating common and rare genetic variation in diverse human populations. *Nature*, 467(7311):52, 2010.

[66] Christopher C Chang, Carson C Chow, Laurent CAM Tellier, Shashaank Vattikuti, Shaun M Purcell, and James J Lee. Second-generation plink: rising to the challenge of larger and richer datasets. *Gigascience*, 4(1):s13742–015, 2015.

[67] Purcell, S. M. and Chang, C. C. PLINK 2.0. `www.cog-genomics.org/plink/2.0/`, 2018. Accessed: 2020-05-15.

[68] Yukinori Okada, Di Wu, Gosia Trynka, Towfique Raj, Chikashi Terao, Katsunori Ikari, Yuta Kochi, Koichiro Ohmura, Akari Suzuki, Shinji Yoshida, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature*, 506(7488):376–381, 2014.

[69] Heribert Schunkert, Inke R König, Sekar Kathiresan, Muredach P Reilly, Themistocles L Assimes, Hilma Holm, Michael Preuss, Alexandre FR Stewart, Maja Barbalic, Christian Gieger, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nature genetics*, 43(4):333–338, 2011.

[70] Wei Zhou, Jonas B Nielsen, Lars G Fritsche, Rounak Dey, Maiken E Gabrielsen, Brooke N Wolford, Jonathon LeFaive, Peter VandeHaar, Sarah A Gagliano, Aliya Gifford, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nature genetics*, 50(9):1335–1341, 2018.

[71] Fredrick R Schumacher, Ali Amin Al Olama, Sonja I Berndt, Sara Benlloch, Mahbubl Ahmed, Edward J Saunders, Tokhir Dadaev, Daniel Leongamornlert, Ezequiel Anokian, Clara Cieza-Borrella, et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nature genetics*, 50(7):928–936, 2018.

[72] Tracy A O'Mara, Dylan M Glubb, Frederic Amant, Daniela Annibali, Katie Ashton, John Attia, Paul L Auer, Matthias W Beckmann, Amanda Black, Manjeet K Bolla, et al. Identification of nine new susceptibility loci for endometrial cancer. *Nature communications*, 9(1):1–12, 2018.

[73] Jie Zheng, A Mesut Erzurumluoglu, Benjamin L Elsworth, John P Kemp, Laurence Howe, Philip C Haycock, Gibran Hemani, Katherine Tansey, Charles Laurin, Beate St Pourcain, et al. Ld hub: a centralized database and web interface to perform ld score regression that maximizes the potential of summary level gwas data for snp heritability and genetic correlation analysis. *Bioinformatics*, 33(2):272–279, 2017.

[74] John Trowsdale and Julian C Knight. Major histocompatibility complex genomics and human disease. *Annual review of genomics and human genetics*, 14:301–323, 2013.

[75] Vasiliki Matzaraki, Vinod Kumar, Cisca Wijmenga, and Alexandra Zhernakova. The mhc locus and genetic susceptibility to autoimmune and infectious diseases. *Genome biology*, 18(1):76, 2017.

[76] Shaun Purcell, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel AR Ferreira, David Bender, Julian Maller, Pamela Sklar, Paul IW De Bakker, Mark J Daly, et al. Plink: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*, 81(3):559–575, 2007.

[77] Dietrich Rothenbacher, Hermann Brenner, Albrecht Hoffmeister, Thomas Mertens, Kenneth Persson, and Wolfgang Koenig. Relationship between infectious burden, systemic inflammatory response, and risk of stable coronary artery disease: role of confounding and reference group. *Atherosclerosis*, 170(2):339–345, 2003.

[78] Debashree Ray and Nilanjan Chatterjee. A powerful method for pleiotropic analysis under composite null hypothesis identifies novel shared loci between type 2 diabetes and prostate cancer. *PLoS genetics*, 16(12):e1009218, 2020.

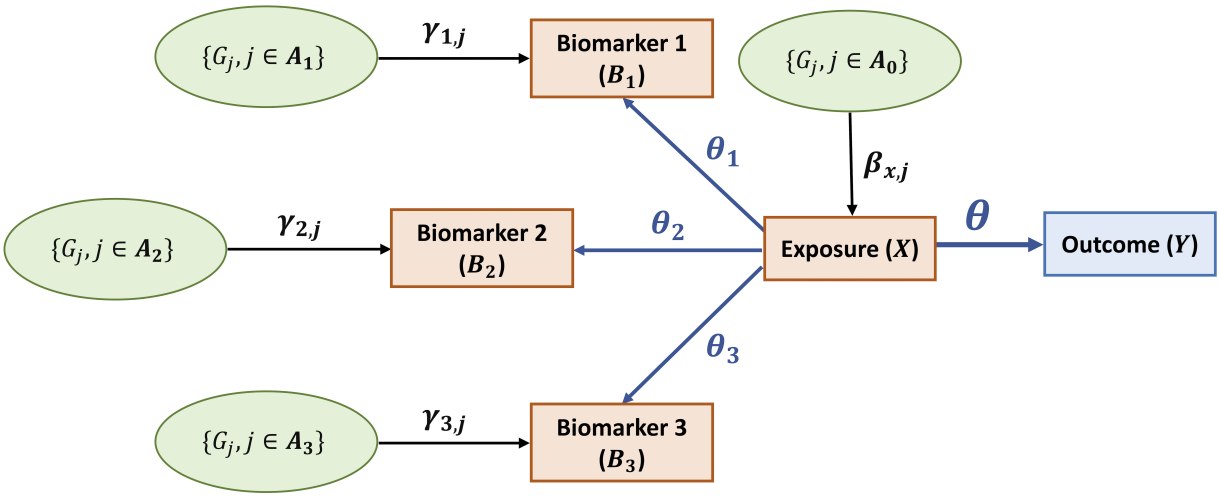

Figure 1. Causal paths between the SNPs (G_js), the latent exposure of interest (X), the biomarkers co-regulated by the latent exposure (B_ks) and the outcome (Y). A_0 represents the set of indexes for the SNPs that are directly associated with X with effect sizes $\beta_{x,j}$s, and A_k represents the set of indexes for the SNPs that are not associated with X but directly associated with B_k with effect sizes $\gamma_{k,j}$s, $k = 1, \ldots, K$. A_0 and A_ks may have overlaps with each other, but the effects of one SNP on different traits or the outcome are independent. The number of observable traits is set to $K = 3$ for illustration.

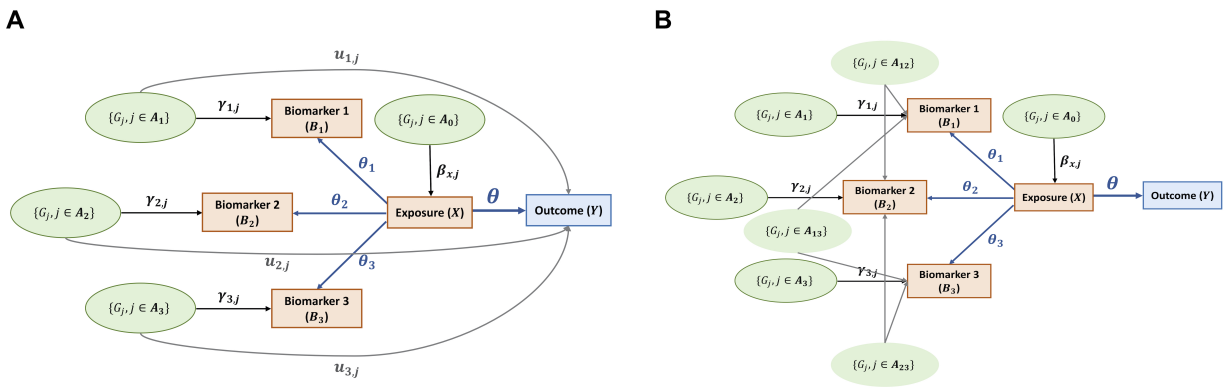

Figure 2. Causal paths between the SNPs (G_js), the latent exposure of interest (X), the biomarkers co-regulated by the latent exposure (B_ks) and the outcome (Y) given potential pleiotropy. (A) Pleiotropic effects across B_ks and Y: some SNPs have correlated associations with Y ($u_{k,j}$s) and at least one of the B_ks ($\gamma_{k,j}$s). (B) Pleiotropic effects across B_ks: some SNPs have correlated direct effects ($\gamma_{k,j}$s) across multiple B_ks.

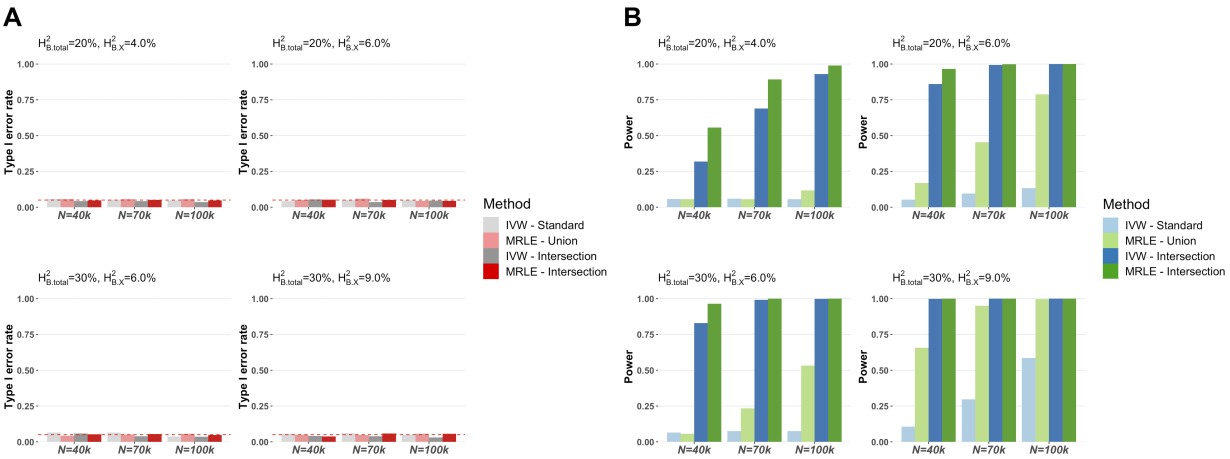

Figure 3. Simulation results assuming a total of $K = 6$ biomarkers based on 1000 simulations per setting. $H^2_{\text{B.total}}$ and $H^2_{\text{B.X}}$ denote respectively the total genetic heritability of each biomarker and the heritability of each biomarker explained by the latent exposure. IVs are defined as either the SNPs associated with at least one biomarker ("IVW-Standard" and "MRLE-Union", $\alpha = 5 \times 10^{-8}$) or the SNPs associated with at least two biomarkers ("IVW-Intersection" and "MRLE-Intersection", $\alpha = 5 \times 10^{-4}$). (A) Barplots showing the empirical type I error rate ($\theta = 0$). (B) Barplots showing the empirical power ($\theta = 0.05$).

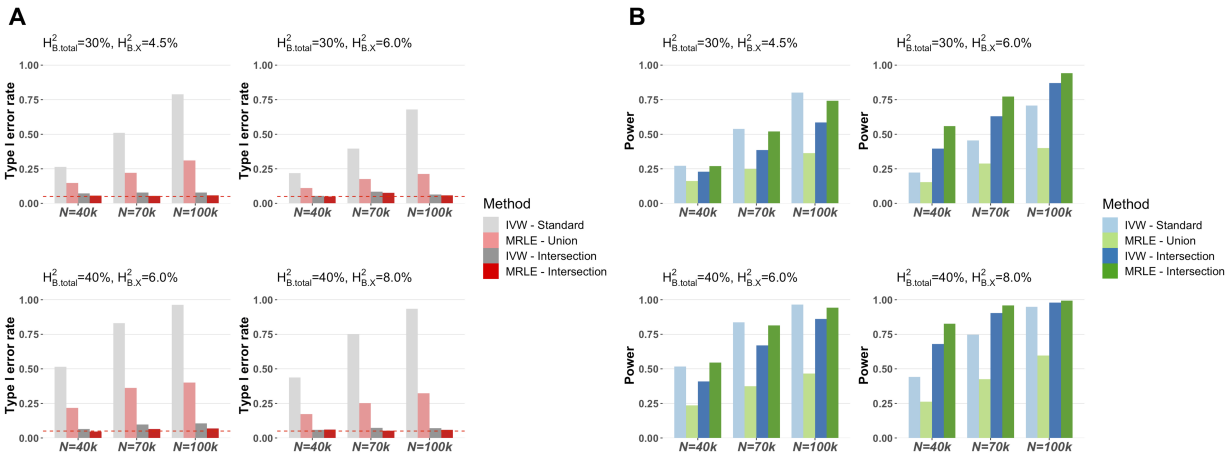

Figure 4. Simulation results assuming a total of $K = 6$ biomarkers and that there are SNPs that have correlated pleiotropic effects across the biomarkers and the outcome. $H^2_{\text{B.total}}$ and $H^2_{\text{B.X}}$ denote respectively the total genetic heritability of each biomarker and the heritability of each biomarker explained by the latent exposure. IVs are defined as either the SNPs associated with at least one biomarker ("IVW-Standard" and "MRLE-Union", $\alpha = 5 \times 10^{-8}$) or the SNPs associated with at least two biomarkers ("IVW-Intersection" and "MRLE-Intersection", $\alpha = 5 \times 10^{-4}$). (A) Barplots showing the empirical type I error rate ($\theta = 0$). (B) Barplots showing the empirical power ($\theta = 0.05$).

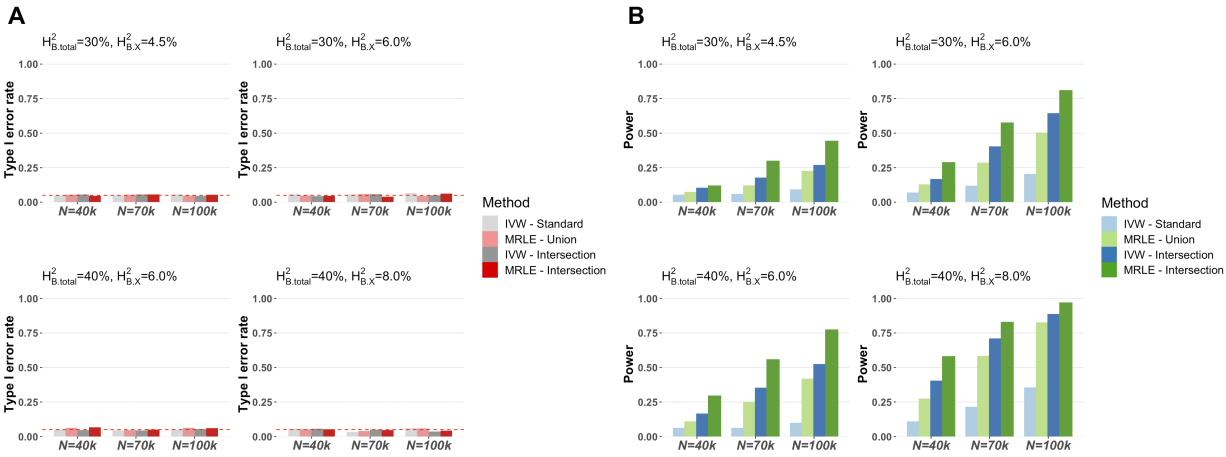

Figure 5. Simulation results assuming a total of $K = 6$ biomarkers and correlated pleiotropic effects across biomarkers. $H^2_{\text{B.total}}$ and $H^2_{\text{B.X}}$ denote respectively the total genetic heritability of each biomarker and the heritability of each biomarker explained by the latent exposure. IVs are defined as either the SNPs associated with at least one biomarker ("IVW-Standard" and "MRLE-Union", $\alpha = 5 \times 10^{-8}$) or the SNPs associated with at least two biomarkers ("IVW-Intersection" and "MRLE-Intersection", $\alpha = 5 \times 10^{-5}$). (A) Barplots showing the empirical type I error rate ($\theta = 0$). (B) Barplots showing the empirical power ($\theta = 0.05$).

Table 1. Causal effect of chronic inflammation on the risk of various diseases. The IVs used in MRLE are the SNPs associated with at least two of the inflammatory biomarkers. The IVs used in each single-biomarker IVW test are the SNPs associated with the biomarker and at least one other biomarker. Significance threshold for IV selection is set to $\alpha = 5 \times 10^{-6}$ for CRP and $\alpha = 10^{-3}$ for the other biomarkers. Bold font indicates significant conclusion. Significance thresholds for IVW tests are set to $1 - (1 - \alpha_0)^{1/K}$, i.e. 0.0102 for CAD, RA and PCa, and 0.0127 for CRC and EC, to control the family-wise error rate at 0.05. CRP were excluded from the tests on CRC and EC due to overlapping individuals in GWAS.

Outcome	Method	Number of IVs	Effect	P-value
RA	IVW - IL-6	22	−	$\mathbf{4.5 \times 10^{-264}}$
	IVW - IL-8	34	+	$\mathbf{6.3 \times 10^{-235}}$
	IVW - TNF-α	30	+	$\mathbf{< 5.0 \times 10^{-324}}$
	IVW - MCP-1	20	−	$\mathbf{9.5 \times 10^{-197}}$
	IVW - CRP	13	+	$\mathbf{1.0 \times 10^{-277}}$
	MRLE	58	+	$\mathbf{5.1 \times 10^{-27}}$
CAD	IVW - IL-6	29	−	0.928
	IVW - IL-8	37	+	0.341
	IVW - TNF-α	35	+	0.243
	IVW - MCP-1	21	−	0.951
	IVW - CRP	15	+	0.097
	MRLE	67	+	**0.012**
CRC	IVW - IL-6	26	+	0.322
	IVW - IL-8	34	+	0.082
	IVW - TNF-α	35	+	0.014
	IVW - MCP-1	13	−	0.408
	MRLE	53	+	**0.011**
PCa	IVW - IL-6	29	−	0.294
	IVW - IL-8	35	+	0.724
	IVW - TNF-α	35	+	0.642
	IVW - MCP-1	20	−	0.229
	IVW - CRP	14	+	0.428
	MRLE	65	+	0.550
EC	IVW - IL-6	26	−	0.325
	IVW - IL-8	34	+	0.541
	IVW - TNF-α	35	+	0.805
	IVW - MCP-1	13	−	0.095
	MRLE	58	−	0.069