Integrating multi-omics summary data using a Mendelian randomization framework

Chong Jin, Brian Lee, Li Shen, Qi Long, for the Alzheimer’s Disease Neuroimaging Initiative and for the Alzheimer’s Disease Metabolomics Consortium

Abstract

Mendelian randomization is a versatile tool to identify the possible causal relationship between an omics biomarker and disease outcome using genetic variants as instrumental variables. A key theme is the prioritization of genes whose omics readouts can be used as predictors of the disease outcome through analyzing GWAS and QTL summary data. However, there is a dearth of study of the best practice in probing the effects of multiple omics biomarkers annotated to the same gene of interest. To bridge this gap, we propose powerful combination tests that integrate multiple correlated P-values without assuming the dependence structure between the exposures. Our extensive simulation experiments demonstrate the superiority of our proposed approach compared with existing methods that are adapted to the setting of our interest. The top hits of the analyses of multi-omics Alzheimer’s disease datasets include genes ABCA7 and ATP1B1.

Keywords: Mendelian randomization, P-value combination, multi-omics data, GWAS, QTL

Introduction

The importance of Mendelian randomization (MR) has been widely recognized [1], particularly in observational studies where investigators probe the effect of an exposure, i.e. a risk factor, on an outcome using heredity as a natural experiment. There is a large body of literature on MR methods particularly for settings with only one molecular exposure [2–5]. However, there has been very limited work and lacks a systematic investigation and consensus on the optimal MR approach(s) when the exposures consist of multiple omics biomarkers such as transcriptomics, proteomics and metabolomics data (Figure 1A). In this work, we seek to evaluate different strategies, some of which are not proposed before under the setting, to integrate P-values when inferring the causal effect (or pleiotropic association) of multiple omics biomarkers on the outcome using an MR framework.

A valid MR investigation rests upon a series of assumptions [6]. These assumptions are as follows.

(i) The genetic variant is associated with the biomarker.
(ii) The genetic variant cannot be associated with any confounder that lies within the biomarker-outcome relationship.
(iii) The genetic variant cannot be associated with the outcome through any pathway other than through the biomarker in question.

The first assumption is fairly easy to satisfy in MR studies. When we employ genetic variants as instrumental variables (IVs) to probe whether a biomarker has a causal effect on an outcome such as disease status [6], we choose genetic variants highly associated with the risk factor so they quality as strong IVs. In contrast, the second and third assumptions are harder to verify due to special challenges in MR: the SNPs are often in linkage disequilibrium (LD) resulting in nebulous interpretations of the association(s) and a large number of candidate exposures may lead to widespread pleiotropy. Thus, our findings of MR analysis need to be interpreted as ‘pleiotropic association(s)’ [7].

To provide a panoramic view of how exposures jointly affect the outcome, instead of having one exposure, we aggregate multiple exposures (Figure 1A). Common examples include different omics biomarkers such as transcriptomics, proteomics and metabolomics [8], multiple tissues [9] and the expression of multiple genes in physical proximity [10]. In this article, we mainly...
explore aggregating the effects of multiple omics biomarkers, though the model can be readily extended to other situations.

Inspired by how multiple traits in multivariate genome-wide association study (GWAS) are modeled [11], we lay out three strategies for modeling multiple exposures in MR: transformation-based, regression-based and combination tests.

(i) Transformation-based tests require a transformation of the measurement of exposures into a single new synthetic exposure through principal component analysis (PCA) or other dimensionality reduction techniques [12, 13]. Note that as opposed to the other two summary data-based approaches, individual-level data access is necessary to enable the transformation.

(ii) Regression models constitute the backbones of MR in practice. If we ignore the causal relationships between the omics biomarkers, then multivariable MR is the appropriate model since the SNPs have direct effects on the biomarkers [14] (Figure 1B).

(iii) Combination tests circumvent the necessity of modeling the relationship between different exposures, which makes the investigation easier under the perspective of meta-analysis of summary data where the relationship between different exposures is largely unknown (Figure 1C).

We focus the attention on combination tests, which have not been applied to the MR setting. The merit of our novel framework is that it can be applied when the structure between multiple exposures is unknown ahead of time or there is no prior knowledge about the structure in a multi-omics study.

In what follows, we first present an overview of the methods and then assess and compare these methods through simulation experiments and analyses of multi-omics datasets of the study of Alzheimer’s disease.

Methods
Existing methods
We begin by describing several existing methods as building blocks toward conducting MR analyses based on multi-omics summary data.

Inverse variance-weighted average method
Inverse variance-weighted average method (IVW) is one of the most popular summary-data-based MR methods. While individual-level genetic data are usually considered sensitive and protected, GWAS and quantitative trait loci (QTL) summary statistics are easier to obtain. These summary statistics can then be used to reach a larger sample size and a higher power in a meta-analysis. Here, we assume that the GWAS and QTL summary statistics are generated from different individuals (two-sample setting) and we have effect size estimates, standard error estimates, statistics, and \( P \)-values for SNPs \( j = 1 \ldots p \) as IVs. In our setting, the \( p \) SNPs are at most 1M base pairs apart from the gene body so that they count as cis-QTLs of the gene.

When the genetic variants are uncorrelated, the IVW estimate of the causal effect size of the outcome on the exposure \( k \) is [6]

\[
\hat{\theta}_{IVW,k} = \frac{\sum_{j=1}^{p} \hat{\beta}_{Xkj} \hat{\beta}_{Yj} \hat{\sigma}_{Yj}^{-2}}{\sum_{j=1}^{p} \hat{\beta}_{Yj}^{-2}}.
\]

where the standard error of the estimate is

\[
se(\hat{\theta}_{IVW,k}) = \sqrt{\frac{1}{\sum_{j=1}^{p} \hat{\beta}_{Yj}^{-2}}}.
\]

Here, \( \hat{\beta}_{Xkj} \) is the regression coefficient of the exposure \( k \) on the SNP \( j \) in the form of QTL and \( \hat{\beta}_{Yj} \) is the regression coefficient of the outcome on the SNP \( j \) in the form of GWAS. \( \hat{\sigma}_{Yj}^{-2} \) is the standard error of the regression coefficient of the outcome on the SNP \( j \).

The omics biomarker \( k \) can stand for transcriptomics, proteomics, metabolomics or other omics readouts. This allows us to combine such \( P \)-values using a method we choose (i.e. Fisher’s combination test, Cauchy combination test, etc.).

Generalized least squares method
The generalized least squares method (GLS) builds upon IVW and accounts for remaining LD within SNPs after LD clumping [5].
having a coefficient of the outcome on SNP j. The notations are the same as they appear in IVW

\[ \hat{\theta}_{k} \approx (\hat{\beta}_{j}^{2} \Omega^{-1} \hat{\beta}_{j}^{2})^{-1}(\hat{\beta}_{j}^{2} \Omega^{-1} \hat{\beta}_{j}^{2})^{-1}, \]

where the variance of the QQTL is estimated using independent samples, we have an approximate test statistic with one degree of freedom

\[ P \approx \frac{\hat{\theta}_{k}^{2}}{\hat{\beta}_{j}^{2} \Omega^{-1} \hat{\beta}_{j}^{2}}. \]

and \( r_{ij} \)'s are the entries in the unsquared LD matrix R.

**Summary data-based MR**

The notable method Summary data-based MR (SMR) was introduced in [7]. The major distinction compared with IVW is that it uses a second-order weight and that information from only one SNP is needed.

Under the assumption that the GWAS z-value \( z_{k} \) and QTL z-value \( z_{xk} \) are estimated using independent samples, we have an approximate test statistic with one degree of freedom

\[ T_{SMR,k} = \frac{z_{k}^{2} z_{xk}^{2}}{z_{k}^{2} + z_{xk}^{2}}. \]

To extend SMR statistics to multiple exposures, we separate the calculation of SMR statistics (one SNP and one omic biomarker) [7] and then combines the P-values. In the method labeled as ‘allSNPs’, P-values of SMR from all combinations of SNPs and omics biomarkers are aggregated, while in the method labeled as ‘singleSNP’, for each omic biomarker, the SNP with the strongest QTL is chosen to derive a P-value of SMR. These P-values from different biomarkers are aggregated.

**Generalized summary data-based MR**

The generalized summary data-based MR (GSMR) method was introduced in [15] to extend SMR by introducing multiple genetic variants that can be correlated to increase the power.

Let

\[ \hat{\theta}_{k} = (\hat{\theta}_{k1}, \hat{\theta}_{k2}, \ldots, \hat{\theta}_{km})^{\top} = \left( \frac{\hat{\beta}_{j1}}{\hat{\beta}_{xk1}}, \frac{\hat{\beta}_{j2}}{\hat{\beta}_{xk2}}, \ldots, \frac{\hat{\beta}_{jm}}{\hat{\beta}_{xkm}} \right)^{\top}, \]

\[ \text{var} (\hat{\theta}_{k}) = \frac{\text{var}(\hat{\beta}_{j})}{\hat{\beta}_{xk}^{2}} \left[ \frac{\text{var}(\hat{\beta}_{xk})}{\hat{\beta}_{xk}^{4}} + \frac{\text{var}(\hat{\beta}_{xk})}{\hat{\beta}_{xk}^{2}} \right], \]

\[ \text{cov} (\hat{\theta}_{k1}, \hat{\theta}_{k2}) = \frac{r_{ij}\sqrt{\text{var}(\hat{\beta}_{j})} \text{var}(\hat{\beta}_{k})}{\hat{\beta}_{xk}^{2}}, \]

\[ V_{k} = (\text{cov}(\hat{\theta}_{k1}, \hat{\theta}_{k2}))_{p \times p}. \]

Here, \( \hat{\beta}_{xk} \) is the regression coefficient of the exposure k on the SNP j in the form of QTL and \( \hat{\beta}_{xk} \) is the regression coefficient of the outcome on the SNP j in the form of GWAS. \( \text{var}(\hat{\beta}_{j}) \) is the variance matrix for the effect size estimates derived from multiple genetic variants.

The test statistic, \( T_{GSMR,k} \), follows a \( \chi^{2} \) distribution with 1 degree of freedom under the null. Note that GSMR needs multiple strong—and weakly correlated—IVs to work. As is implemented in the GSMR R package, the default filtering conditions on the SNPs are \( r_{ij}^{2} < 0.95 \) and \( \chi^{2} > 10 \) for the SNPs, and after the filtering, at least 10 SNPs are recommended. Otherwise, the method often cannot work due to the covariance matrix being not invertible.

**Multivariable MR**

When a genetic variant is causally linked to multiple related exposures, multivariable MR may be adopted as the primary analysis strategy [2]. Multivariable MR under one sample and two sample settings is extensively discussed in [16]. In multivariable MR, the effect sizes \( \theta_{k} \)'s are coefficients in the regression model

\[ \hat{\beta}_{j} = \sum_{k=1}^{m} \theta_{k} \hat{\beta}_{xk} + \epsilon, \text{ weights } = \hat{\beta}_{xk}^{2}. \]

where \( \hat{\beta}_{j} \) is the regression coefficient of the outcome on SNP j and \( \hat{\beta}_{xk} \) is the regression coefficient of exposure k on the SNP j in \( \{1 \ldots p\} \).

To perform a joint test whether there is any exposure k having a pleiotropic effect on the outcome using multivariable MR, we test the null hypothesis of

\[ \theta_{1} = \theta_{2} = \ldots = \theta_{k} = \ldots = \theta_{m} = 0. \]

This results in a test with m degrees of freedom.

Multivariable MR can be viewed as an extension of the IVW method (and the GLS method in presence of LD) [17]. More IVs (variants) are needed than exposures. Otherwise, the fitted values of different exposures by plugging in the IVs will not constitute a full rank matrix. While multivariable MR generally requires uncorrelated variants, it can be extended to incorporate correlated variants using GLS [5, 10].

The interpretation of the effect size estimate differs from the interpretation when we only have one exposure. Multivariable MR quantifies the direct effect of an exposure on the outcome as opposed to the total effect of an exposure on the outcome using conventional univariable MR [16].

**Proposed method: combining omics biomarkers annotated to one gene**

We propose combination tests, which have not been applied to the multi-omics MR setting before, to investigate the causal effects of multiple omics biomarkers on an outcome. When conducting the combination tests, we first test the effect of each omics biomarker on the outcome separately using an MR method we covered above and then combine the P-values using the methods detailed below. The combinations of strategies are listed in Figure 2A.
Figure 2. (A) Overview of MR methods and combination methods illustrated in our benchmark. Note that Fisher_gamma is only implemented for GLS, and MultivariableMR can be combined with either IVW and GLS. More results of Fisher_chisq and MinP-related methods can be found in the online supporting data repository. (B) P-value combination methods when combining two P-values. The first P-value is plotted on the x-axis and the second P-value is plotted along the panels. The function in the Cauchy combination method (grey solid line) is very similar to the function in harmonic mean P-value (brown dashed line) albeit the latter one is a bit more conservative; both methods are valid regardless of the correlation between P-values. The Fisher combination method using chi-squared approximation 'Fisher_chisq' (beige dotted line), more liberal than others, was originally proposed for independent tests and can be adapted to 'Fisher_gamma' where the covariance structure of the z-values underlying the P-values is known (black dot-dashed line; correlation between two z-values is 0.5 as an example).

- Cauchy combination test \([18-20]\): The method calculates a P-value for each omic biomarker and then combines the P-values using Cauchy combination function and Cauchy distribution approximation

\[
T_{\text{Cauchy}} = \sum_{k=1}^{m} w_k \tan((0.5 - p_k)\pi),
\]

where \(w_k\)'s are weights determined externally. A natural choice of equal weights \(w_k = 1/m\) is used here.

- Regardless of the correlation structure between P-values \(p_k\), \(T_{\text{Cauchy}}\) can be approximated using Cauchy distribution, and the P-value of the Cauchy combination test is

\[
0.5 - \arctan \left( \frac{T_{\text{Cauchy}}}{\sum_{k=1}^{m} w_k} \right) / \pi.
\]
Note that when $T_{\text{Cauchy}}$ is very small, we may use other approximations during computation for numerical issues.

- Harmonic mean $P$-value (HMP): HMP is a $P$-value combination method originated from Bayesian model averaging offered as an alternative to Fisher’s method when the $P$-values are correlated [21]

$$T_{\text{HMP}} = \sum_{k=1}^{m} \frac{u_k}{\sum_{k=1}^{m} u_k / p_k}$$

The HMP test statistics are then compared with a Stable distribution with tail index $\lambda = \beta = 1$ to derive the $P$-value.

- Minimum $P$-value (MinP): The statistic is

$$T_{\text{MinP}} = \min_{k \in \{1, \ldots, m\}} p_k,$$

and when only summary data are available, we ignore the correlation between $P$-values and take the statistic $T_{\text{MinP}}$ as the $P$-value.

- Fisher_chisq: This method calculates a $P$-value for each omic biomarker and then combines the $P$-values using Fisher combination function and the chi-squared approximation

$$T_{\text{Fisher}} = -2 \sum_{k=1}^{m} \ln(p_k).$$

The Fisher combination statistics are compared with a chi-squared distribution of $2m$ degrees of freedom when the $P$-values are independent.

- When the $P$-values are not independent and we know the association between the $P$-values (can be further approximated using the correlation between the measurements of omics biomarkers), a gamma distribution approximation would be more appropriate, which is discussed below [11, 22, 23].

- Fisher_gamma: We propose the method that combines $P$-values from the GLS method for each omic biomarker. The combination of $P$-values involves the Fisher combination function and gamma distribution approximation of the null hypothesis given the correlation between the summary $z$-statistics as described in Yang et al. [11]. This concludes our proposed method of using Fisher combination function and gamma distribution approximation of the null hypothesis to combine $P$-values representing multiple omics biomarkers obtained from GLS.

Cauchy combination method and HMP, both being valid regardless of correlation between $P$-values, give similar combined $P$-values with Cauchy combination method having a slight edge. While the small $P$-value-producing Fisher combination method is invalid when there is correlation between the two $z$-values as results from Mendelian randomization, Fisher combination methods with a known correlation can strike a trade-off between retaining a small combined $P$-value and relaxing the correlation structure assumptions (Figure 2B).

**Simulation**

**Simulation setup**

To assess different MR frameworks when combining summary data from multiple omics biomarkers, we perform a comprehensive study of them under different simulation setups. In each simulation setup, we first simulate individual-level SNPs, multi-omics data and outcome labels. Then, the summary statistics are computed by fitting regression models and are fed to the computational methods that provide a $P$-value of whether there is any pleiotropic association between the omic biomarker (exposure) and the outcome. The procedure is repeated 10 000 times and the empirical power and type I error are recorded for each method in each simulation setup.

The total number of subjects in the simulation was chosen to be 500 to represent a typical sample size in a multi-omics study (i.e. the ADNI data used in the real data analysis example). While straightforward to simulate 500 subjects when the outcome is continuous, it requires down-sampling to get the desired number of case and control subjects when the outcome is binary.
To generate individual-level data of 250 case and 250 control subjects with a targeted level of subject overlap between QTL and GWAS datasets, we first simulate a population with a much larger number of samples (up to 5000) and generate the outcome label Y. We pick 250 case and 250 control subjects from the simulated population to use as samples to generate the GWAS summary statistics. We also pick 250 control subjects with a pre-specified level of overlap (0, 0.5, 1) with the aforementioned 250 control subjects from the simulated population to use as samples to generate the QTL summary statistics.

We assume the SNP effects on omics biomarkers X1, X2, and X3 are constant. We have additional simulations where the SNP effects on omics biomarkers follow an i.i.d. beta distribution. For each individual, the data are generated as below.

\[ Z_j \sim \text{Binom}(2, 0.3) \text{ for } j = 1, \ldots, p, \]
\[ r^2_{ij} = 0, 0.01, \text{ or } 0.2 \text{ for } i \neq j, \]
\[ X_1 = \frac{\alpha}{p} \sum_{j=1}^{p} Z_j + U + \epsilon_{X1}, \]
\[ X_2 = 2X_1 - U + \epsilon_{X2}, \]
\[ X_3 = -0.5X_2 + U + \epsilon_{X3}, \]
\[ Y_b \sim \text{Binom}(1, r), \text{ where } \logit(r) = -2 + \beta_X X_3 - U, \]
\[ Y_C = \beta_X X_3 - U + \epsilon_Y, \]
\[ U \sim N(0, 1), \epsilon_{X1} \sim N(0, 1), \epsilon_Y \sim N(0, 1) \text{ independently.} \]

Under horizontal pleiotropy, let Yc be the continuous outcome and Yb be the binary outcome. For each individual, the data are generated as below.

\[ Z_j \sim \text{Binom}(2, 0.3) \text{ for } j = 1, \ldots, p, \]
\[ r^2_{ij} = 0, 0.01, \text{ or } 0.2 \text{ for } i \neq j, \]
\[ X_1 = \frac{\alpha}{p} \sum_{j=1}^{p} Z_j + U + \epsilon_{X1}, \]
\[ X_2 = -\frac{\alpha}{p} \sum_{j=1}^{p} Y_j - U + \epsilon_{X2}, \]
\[ X_3 = \frac{\alpha}{p} \sum_{j=1}^{p} Y_j - U + \epsilon_{X3}, \]
\[ Y_b \sim \text{Binom}(1, r), \text{ where } \logit(r) = -2 + \beta_X \sum_{k=1}^{3} X_k + U, \]
\[ Y_C = \beta_X \sum_{k=1}^{3} X_k + U + \epsilon_Y, \]
\[ U \sim N(0, 1), \epsilon_{X1} \sim N(0, 1), \epsilon_Y \sim N(0, 1) \text{ independently.} \]

Then, we generate the summary data, which are the input of the MR analyses. For the association between exposure Xk and genetic variant Zj, we fit the regression model to get effect size estimate \( \hat{\beta}_{Xk} \) and standard error estimate \( \hat{\sigma}_{Xk} \) for an omics biomarker k = 1, 2, 3 and genetic variant j

\[ X_k \sim \beta_X x_j + \beta_Z Z_j, \]

For the association between outcome Y and genetic variant Zj, when we have binary outcome Yb, we fit the logistic regression model to get effect size estimate \( \hat{\beta}_{Yj} \) and standard error estimate \( \hat{\sigma}_{Yj} \)

\[ \logit(Y_b) \sim \beta_{Yj} + \beta_{Yj} Z_j, \]

and when we have continuous outcome Yc, we fit the regression model to get effect size estimate \( \hat{\beta}_{Yj} \) and standard error estimate \( \hat{\sigma}_{Yj} \)

\[ Y_C \sim \beta_Y + \beta_{Yj} Z_j, \]

Below is a list of simulation settings.

(i) Pleiotropy: horizontal or vertical (Figure 5);
(ii) Number of SNPs used as IVs: 5 or 20;
(iii) Overlap between samples involved in the GWAS studies and the QTL studies: 0 (two-sample), 0.5 (half of the samples overlap) or 1 (one-sample);
(iv) Use same samples or independent samples to calculate QTLs belonging to multiple omics biomarkers under a two-sample setting;
(v) Outcome: continuous (Yc) or binary (Yb). When the outcome is binary, we also need to decide whether we will only use control samples to estimate the QTLs;
(vi) LD between different SNPs \( r^2 \): 0, 0.01 or 0.2;
(vii) Strength of IV \( \alpha \), which is proportional to the association between IVs and exposures: 0.5, 1 or 2;
(viii) Effect size of the pleiotropic association between exposure and outcome \( \beta_X \): 0 or 0.1.

We set \( \beta_X = 0.1 \) (positive casual effect) to evaluate the power and set \( \beta_X = 0 \) (null casual effect) to evaluate the type I error. We set the IV strength \( \alpha \) to be either 0.5, 1 or 2.

Comparison of type I error across methods

In the subsequent simulations, we eliminate the methods that fail to control type I error (methods with a shaded background in Figures 2A and 3) and identify the ones whose powers stand out from the pack after another round of competition in Figure 4.

First, we verify whether GLSs help us adjust for LD between SNPs used as IVs after pre-emptively striking out some scenarios where the excessive participants overlap between GWAS and QTL studies could bring bias.

Figure 3 illustrates that when the LD is not accounted for in the MR analysis, the type I error tends to increase sharply especially when there is a larger number of SNPs (20 vs. 5), even the true LD between SNPs is only 0.01. Given the difficulty of accurately estimating the LD matrix, the most prudent course of action is to select SNPs known to be uncorrelated to use as IVs, which is not practical since the association between an SNP and an exposure is often local and we may fail to identify a considerable number of independent SNPs as valid IVs. The comparison of methods stresses that a valid method needs to account for LD between SNPs using GLSs under regression-based models.
Figure 3. A comparison of type I error rates across combinations of MR methods and combination methods. A method is disqualified if its type I error rate exceeds the nominal significance level of 0.05 (red horizontal line). This can be due to a failure in correcting for LD between SNPs (labeled in light green) or poor approximation of null distribution when the underlying \( P \)-values are correlated, including Fisher combination function using chi-squared distribution (Fisher_chisq) and minimum \( P \)-value (MinP) (labeled in light blue). Each observation in the boxplot represents the type I error rate of a simulation setting with 10,000 replications. Simulation settings are included if either (1) outcome is binary and only control samples are used when estimating QTL or (2) when there is no overlap between samples in the QTL and the GWAS datasets.

Then, we verify whether there is an inflated type I error rate associated with the use of several \( P \)-value combination methods. The \( P \)-value combination methods need to be valid in the presence of correlation between \( P \)-values. Observed from left to right of the \( x \)-axis labels of Figure 3, Cauchy, HMP and Fisher combination method with a gamma distribution approximation (Fisher_gamma) can control type I error. This is expected because both the HMP and the Cauchy combination method work regardless of the correlation between \( P \)-values. Meanwhile, Fisher combination method with a chi-squared distribution approximation (Fisher_chisq), which does not adjust for correlation between \( P \)-values, and the minimum \( P \)-value method (MinP) both fail to control type I error.

Resolution to sample overlap between GWAS and QTL studies

We discuss what level of sample overlap between GWAS and QTL studies is appropriate to be incorporated in an MR analysis. It has been reported in the literature that the causal effect size estimate in MR will be biased towards the OLS estimate, which is the direction of there being no effect when the samples in the GWAS and the samples in QTL studies have no overlap since they are independent (‘two-sample’). When there is considerable overlap between samples involved in the GWAS and QTL studies (‘one-sample’), however, the MR causal effect size estimate will be biased towards the direction of there being an effect, which is usually nonzero under the presence of the confounders [28].

Our findings mostly align with what is reported in the literature with some subtleties. In Figure 4, when the overlap is 1, which means the samples underlying the GWAS and QTL studies are the same, there is an inflated type I error for the methods that rely on a model-based approach to integrate the information from multiple SNPs, and the symptom is more prominent when the number of SNPs is 20 instead of 5. The methods based on combining \( P \)-values arising from one-SNP-based SMR methods can control type I errors surprisingly well considering that SMR requires two samples. Other than the obvious corollary that we need to pick a GWAS dataset without sample overlap with QTL datasets, we can also bring type I error under control when the outcome is binary and only control samples are considered when estimating the QTLs.

Comparison of power across methods

After we narrow down the candidate methods to those will not bring inflated type I errors, we pit them against each other and figure out whose power is the highest. In Figure 5, we find that GLS_Fisher_gamma and GLS_Cauchy are the most powerful ones throughout different scenarios, and GSMR_Cauchy is about as powerful when the IVs are strong but is less powerful with weak IVs. We also confirm that the methods being examined will not produce inflated type I error rates.
Figure 4. Type I error rate of combinations of MR methods and combination methods when the QTL and the GWAS datasets are simulated with various degrees of overlap between samples (rows of the grid) and outcome types being continuous or binary (columns of the grid). Simulation settings that have inflated type I errors are labeled in light purple. Each observation in the boxplot represents the type I error rate under a simulation scenario with 10,000 replications. MR methods that are not able to account for correlation between SNPs and/or correlation between \(P\)-values of MR analyses conducted in different exposures are excluded. The red horizontal line depicts the nominal significance level of 0.05.

Figure 5. Type I error rate (middle column) and power (right column) of selected combinations of MR methods and combination methods whose empirical type I error rates do not exceed the nominal significance level. The first row of panels represents the results under horizontal pleiotropy and the second row of panels represents the results under vertical pleiotropy. The number of SNPs in the simulation is 5 (only 3 SNPs are depicted in the diagram due to space constraints) and the \(r^2\) between the SNPs is 0.2. In the simulation, there is no overlap between samples used in QTL and GWAS datasets, and both case and control samples are used when estimating QTLs. Each point represents the positive rate under a simulation scenario with 10,000 replications. The red horizontal line depicts the nominal significance level of 0.05.
We verify if our combination method works under different genetic architecture of traits through additional simulations. Similar type I error rates and powers can be observed when

(i) The number of SNPs changes from 5 to 20 (Supporting Figure 1);
(ii) The causal effects of SNPs on omics biomarkers follow i.i.d. beta distributions instead of being constant (Supporting Figures 3, 4, and Web Appendix B);
(iii) The phenotype and the omics biomarkers are simulated using real genotypes from the 1000 Genomes Project (Supporting Figure 6 and Web Appendix C).

These results provide additional supporting evidence for the utility of our proposed framework.

Prioritization of genes linked to Alzheimer’s disease using expression, proteomics and metabolomics biomarkers

Study samples and real datasets

We analyze multi-omics data related to Alzheimer’s disease to assess the performance of different methods on identifying important genes where the omics biomarkers have pleiotropic associations with outcome. Data used in our analysis were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). ADNI was launched in 2003 as a public–private partnership to test whether serial MRI, PET and biological markers can be combined with clinical and neuropsychological assessments to accurately measure the progression of mild cognitive impairment (MCI) and early AD. For up-to-date information, see www.adni-info.org. We use genotype, methylation and metabolomics data derived from members of the ADNI cohort. See Web Appendix A in Supporting Information for how the QTLs of omics data are calculated [29].

Downloading GWAS summary data

The GWAS summary data are obtained from a meta-analysis study [30] of European ancestry individuals. This study integrates the GWAS summary statistics from Kunkle et al. [31] (22,000 cases and 42,000 controls) with the GWAX (GWAS by proxy) summary statistics in UKB (53,000 proxy cases and 378,000 controls). Individuals were denoted as ‘cases’ if either of their or any siblings had AD. The data were downloaded from http://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST90012001-GCST90013000/GCST90012877/.

Steps of real data analysis

We further analyze the real datasets described earlier to assess the performance of the proposed methods.

(i) Prepare gene-level metQTL summary data by mapping the metabolites to genes using KEGG metabolic pathways. If several metabolites in SNP-metabolite associations are mapped to one gene, the metQTL to be retained is the one with the smallest P-value.
(ii) Prepare gene-level pQTL summary data by mapping the UNIPROT AC/ID to Gene name on https://www.uniprot.org/uploadlists/.
(iii) Merging downloaded metQTL and pQTL into the eQTL meta-analysis summary statistics by aligning SNP IDs and gene symbols.
(iv) Select a list of 584 genes with some evidence of association to Alzheimer’s disease from Agora genes at https://agora.acknowledgeportal.org/genes.
(v) Within the Agora genes, select 41 genes where all three omics biomarkers (expression, proteomics and metabolomics) are present.
(vi) Select 27 genes that at least one of eQTL, pQTL and metQTL is significant (P-value < 0.05).
(vii) For each gene, select cis-SNPs (distance between gene and SNP less than 100 kb) and perform LD clumping in each omics biomarker (select SNP with the smallest QTL when possible, \( r^2 < 0.2 \), distance > 100 kb).
(viii) Perform MR analysis for each omics biomarker using the exposure datasets and the outcome datasets.
(ix) Integrate P-values using combination tests.

Results

In this real example of a multi-omics MR analysis, we first decide if we should proceed with a one-sample or a two-sample framework. The former corresponds to using QTL datasets whose participants overlap those in the GWAS dataset, while the latter requires QTL datasets unrelated to the GWAS dataset. The main argument for the two-sample MR framework is that the sample size of the GWAS dataset is several magnitudes larger, enabling more discoveries. Another benefit is that we are not plagued by the over-inflated type I error rate that would arise in a one-sample framework unless the QTL summary data are derived solely from control subjects. Therefore, we aim to use an external GWAS dataset (meta-analysis of AD by proxy [30]) to perform a two-sample MR.

We start from a list of 584 ‘Nominated Targets’ from the Agora website [32], which include genes that are potential targets of Alzheimer’s disease identified through genomic, proteomic and/or metabolomic data analysis. After selecting the genes that are covered by the expression, proteomic and metabolomic assays, we have 41 genes. Next, we stipulate that any genes left would need to have at least one of eQTL, pQTL and metQTL being significant (Bonferroni-corrected P-value < 0.05/41), so the associated SNPs can be selected as IVs, and 27 genes survive this threshold. Then, we perform a targeted analysis using these prominent gene targets.

Fisher combination test using a gamma distribution approximation identifies two significant findings marked using an asterisk * (Figure 6). The panels illustrate three different methods combining MR results on each gene with mapped eQTL, pQTL and metQTL. The horizontal red line represents the cutoff of a significance level of 0.05 after adjusting for multiple comparisons using the Bonferroni method. Focusing on the genes that pass the cutoffs, the gene ABCA7 is the most significant finding when using either the Cauchy combination test or Fisher combination test. There are no other genome-wide significant findings in the Cauchy combination test, while the Fisher combination test (using a Gamma distribution approximation) picks another significant gene in ATP1B1.

Both ABCA7 and ATP1B1 are candidate biomarkers of the Alzheimer’s disease. ABCA7 is a protein-encoding gene that creates ATP-binding cassette sub-family A member 7. This protein is one of the ATP-binding cassette (ABC) transporters, which generally transports molecules across cellular membranes [33]. To date, two large genome-wide association studies (GWAS) [34, 35] have implicated ABCA7 as a susceptibility locus for late-onset Alzheimer’s Disease. Specifically, ABCA7 has been associated with...
memory decline [36], incident MCI [36] and scortical/hippocampal atrophy [37]. These findings were also biologically verified using knock-out mice: deletion of ABCA7 resulted in the increased production of A-beta proteins indicative of AD [38, 39]. The ATP1B1 gene encodes the eponymous protein ATP1B1 (sodium/potassium-transporting ATPase subunit beta-1) [40]. This protein is part of an enzyme that catalyzes the hydrolysis of ATP which is coupled with the exchange of Na\textsuperscript{+} and K\textsuperscript{+} ions across the plasma membrane. The beta subunit this protein contributes to regulates the quantity of sodium pumps transported to the plasma membrane and is therefore essential to neuron function and health [41].

Other combination tests and multivariable MR fail to report any significant findings after a multiple adjustment. Using the GSMR approach with the Cauchy combination test, the smallest \(P\)-value among the 27 tests turns out to be 0.0254, which is not genome-wide significant after Bonferroni adjustment for 27 tests. This agrees with the simulation result in Figure 5: when the IVs are weaker, the GSMR approach is less powerful compared with GLS. The GSMR being less powerful is partly due to mostly only two or three IVs left after the LD clumping, which fell short of the cutoff in the GSMR implementation recommending at least 10 strong IVs. Likewise, the application of multivariable MR requires an even more stringent number of IVs. The test can proceed only if there are at least four IVs left after the LD clumping. As a result, out of 27 genes we tested, only in two genes multivariable MR produces non-NA \(P\)-values, both of which are not significant.

**Discussion**

MR methods can identify pleiotropic association using GWAS and QTL summary data. Pleiotropy happens when an SNP has multiple effects, and it is widespread in MR investigations considering that an SNP can affect more than one omics biomarker. When QTL data from multiple omics biomarkers are available, the MR analyses that focused on one gene can be integrated using \(P\)-value combination methods. Some \(P\)-value combination methods can account for the correlation between the MR test statistics. They include the Cauchy combination test, harmonized mean \(P\)-value (HMP) and Fisher combination function with gamma distribution approximation. The simulations and real data analysis favor Fisher combination function with gamma distribution approximation, especially when the IVs are weak.

We explore how multi-omics data can be integrated into the MR framework. Pleiotropy, which happens when an SNP has multiple effects, is very common in multi-omics studies. Previously, researchers distinguished vertical pleiotropy from horizontal pleiotropy (Figure 5). Vertical pleiotropy can be investigated using multi-step MR by picking exposure and outcome as omics-data types where we want to confirm a causal relationship [42]. On the other hand, multivariable MR suits horizontal pleiotropy better where the independent effects of multiple exposures can be estimated in one model. Our novel approach focuses on combining multiple \(P\)-values after performing MR in...
each -omics biomarker. Such prioritization of genes in a genome-wide analysis achieves an increase in power at the expense of deeper interrogations of relationship between biomarkers.

In our MR experiment investigating GWAS and QTLs, we consider cis-QTLs which are associations between variants and exposure at most one megabase pairs apart. For an exposure, there might not be multiple SNPs with significant cis-QTLs after LD clumping of the candidate SNPs. Considering that a majority of SNPs affect the multi-omic exposures locally, we usually do not have multiple independent SNPs as IVs due to LD, restricting our capability of further interpreting MR results. Therefore, it is usually not practical to tell apart causality from pleiotropy and linkage [7].

Apart from one-sample and two-sample MR, there is another approach known as three-sample MR [43], whereby we can use selection datasets with no overlap in the individuals compared with the exposure datasets. If we use the same datasets both to select the SNPs and to provide statistics to perform MR analysis as shown in the real data analysis example, we may suffer from the ‘winner’s curse’ and generally have a causal effect size estimate biased towards zero. A future direction to our proposed strategy is to adopt a three-sample approach by finding other eQTL, pQTL and metQTL datasets for the MR analysis.

We note that the mediation analysis model attempts to evaluate the joint effect of genotype on outcome [44, 45]. MR is also compared with fine mapping and colocalization [46], both methods working on a single gene region, where fine mapping strives to find causal genetic variants for a trait and colocalization aims to identify variants that are associated with both traits. The concept of MR is unique in that it quantifies the causal relationship from exposure to outcome (if the assumptions are satisfied and the effect is not due to pleiotropy or linkage) and the exposures may serve as targets for intervention [47].

In conclusion, we developed a framework for summary-data-based MR analysis where multiple omics biomarkers can be viewed as multiple exposures, with an emphasis on the combination tests and special handling due to correlated P-values from single-exposure MR tests. While our proposed framework does not assume the dependence structure between the omics biomarkers, we anticipate that it can be complemented by specifically designed tests that can untangle the effect of each omics biomarker in future research.

**Key Points**

- We propose combination tests that aggregate P-values related to a gene after Mendelian randomization (MR) analyses probing the causal effect (or pleiotropic effect) of omics biomarkers on the outcome.
- Both in simulations and a real example, the combination tests are more powerful in gene prioritization than the multivariable MR framework.
- We recommend using Fisher combination test with gamma distribution approximation. Its power is the highest among the compared methods especially when only weak IVs can be selected from the variants.

**Data availability statement**

The source code and supporting data are available through GitHub (https://github.com/lshen/adomics). Processed eQTL data used in the analysis are from the ROS/MAP, MayoRNAseq studies and are available after application through the AMP-AD knowledge portal (https://adknowledgeportal.synapse.org) [48]. Processed pQTL data are from the ROS/MAP project [49]. Methylation data, metabolomics data and genotypes used to generate meQTL and metQTL in this analysis are from the ADNI project and are available after application through the Alzheimer’s Disease Neuroimaging Initiative website (http://adni.loni.usc.edu/).

**Authors’ contributions statement**

Q.L. and L.S. initiated the study. Q.L., L.S. and C.J. designed the study. B.L. and C.J. processed real data. C.J. implemented the methods, benchmarked them on simulations and real data and generated visualizations. C.J., B.L., L.S., and Q.L. wrote, revised and reviewed the manuscript.

**Acknowledgements**

For the genetic and genomic data used in the real data analysis, the results published here are in whole or in part based on data obtained from the AD Knowledge Portal (https://adknowledgeportal.org). For the ROSMAP study, we thank the patients and their families for their selfless donation to further understanding Alzheimer’s disease. Many data and biomaterials were collected from several National Institute on Aging (NIA) and National Alzheimer’s Coordinating Center (NACC) funded sites. Amanda J. Myers, PhD (University of Miami, Department of Psychiatry) prepared the series. The directors, pathologist and technicians involved include: Rush University Medical Center, Rush Alzheimer’s Disease Center (NIH AG10161): David A. Bennett, M.D. Julie A. Schneider, MD, MS, Karen Skish, MS, PA (ASCP)MT, Wayne T Longman. Quality control checks and preparation of the gene expression data was provided by the National Institute on Aging Alzheimer’s Disease Data Storage Site (NIAGADS, U24AG041689) at the University of Pennsylvania. For the MSBB study, these were generated from postmortem brain tissue collected through the Mount Sinai VA Medical Center Brain Bank and were provided by Dr Eric Schadt from Mount Sinai School of Medicine. The Mayo RNAseq study data were led by Dr Nilufer Ertekin-Taner, Mayo Clinic, Jacksonville, FL as part of the multi-PI U01 AG046139 (MPIs Golde, Ertekin-Taner, Younkin, Price). Samples were provided from the following sources: The Mayo Clinic Brain Bank. Study data include samples collected through the Sun Health Research Institute Brain and Body Donation Program of Sun City, Arizona. The ADNI study data were provided by the Rush Alzheimer’s Disease Center, Rush University Medical Center, Chicago. The Mayo Clinic Alzheimers Disease Genetic Studies, led by Dr Nilufer Taner and Dr Steven G. Younkin, Mayo Clinic, Jacksonville, FL. Study data includes samples collected through the Sun Health Research Institute Brain and Body Donation Program of Sun City, Arizona. Brain tissue for the study was obtained from the following brain bank collections: the Mount Sinai NIH Brain and Tissue Repository, the University of Pennsylvania Alzheimer’s Disease Core Center, the University of Pittsburgh NeuroBioBank and Brain and Tissue Repositories, and the NIMH Human Brain Collection Core. CMC Leadership: Panos Roussos, Joseph Buxbaum, Andrew Chess, Shachram Akbarian, Vahram Haroutunian (Icahn School of Medicine at Mount Sinai), Bernie Devlin, David Lewis (University of Pittsburgh), Raquel Gur, Chang-Cyu Hahn (University of Pennsylvania), Enrico Domenici (University of Trento), Mette A. Peters, Solveig Sieberts (Sage
Bionetworks), Thomas Lehner, Stefano Marenco, Barbara K. Lipska (NIMH).

**Funding**

The National Institutes of Health (grant numbers RF1 AG063481, R01 AG071470, U01 AG068057); National Institute on Aging (AG034504, AG041232, P30AG10161, R01AG15819, R01AG17917, R01AG36836, R01AG648015, U01AG64152, P50 AG06574, R01 AG032990, U01 AG046139, R01AG18023, U01AG06576, U01 AG006786, R01 AG025711, R01 AG017216, R01 AG003949); National Alzheimer’s Coordinating Center (NACC, grant U01 AG016976); National Institutes of Health (grants P30AG10161, R01AG15819, R01AG17917, R01AG36042, R01AG36836, U01AG64152, R01AG34374, R01NS7809, U18NS82140, R01AG42210, R01AG39478); Illinois Department of Public Health; National Institute on Aging (grants P50 AG016574, R01 AG032990, U01 AG046139, R01 AG017216, R01 AG003949); National Institute of Neurological Disorders and Stroke (grants R01 NS080820, R01 NS080820); CurePSP Foundation; Mayo Foundation; National Institute of Neurological Disorders and Stroke (U24 NS072026 National Brain and Tissue Resource for Parkinsons Disease and Related Disorders); National Institute on Aging (P30 AG19610 Arizona Alzheimers Disease Core Center); Arizona Department of Health Services (contract 211002, Arizona Alzheimers Research Center); Arizona Biomedical Research Commission (contracts 4001, 0011, 05-901 and 1001 to the Arizona Parkinson’s Disease Consortium); Michael J. Fox Foundation for Parkinsons Research; Illinois Department of Public Health (ROSMAP); Translational Genomics Research Institute (genomic); CurePSP Foundation; Takeda Pharmaceuticals Company Limited; F Hoffmann-La Roche Ltd and National Institute of Health (grants R01MH085542, R01MH093725, P50MH066392, P50MH080405, R01MH097276, R01-MH-075916, P50MH06891, P50MH8045351, R37MH057881, AG02219, AG05138, MH06692, R01MH110921, R01M109677, R01MH109897, U01M103932 and contract HHSN271201300031C through IRP NIMH).

**References**

1. Adam D. The gene-based hack that is revolutionizing epidemiology. Nature 2019;576(7786):196–9.
2. Burgess S, Smith GD, Davies NM, et al. Guidelines for performing Mendelian randomization investigations. Wellcome Open Research 2020;4:186.
3. Yuan Z, Zhu H, Zeng P, et al. Testing and controlling for horizontal pleiotropy with probabilistic Mendelian randomization in transcriptome-wide association studies. Nat Commun 2020;11(1):3861.
4. Zheng J, Baird D, Borges M-C, et al. Recent developments in Mendelian randomization studies. Curr Epidemiol Rep 2017;4(4):330-45.
5. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. Stat Med 2016;35(11):1880-906.
6. Burgess S, Scott RA, Timpson NJ, et al. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. Eur J Epidemiol 2015;30(7):543-52.
7. Zhu Z, Zhang F, Han H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 2016;48(5):481-7.
8. Yang W, Zeng J, Zhang F, et al. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. Nat Commun 2018;9(1):918.
9. Zhou D, Jiang Y, Zhong X, et al. A unified framework for joint-tissue transcriptome-wide association and Mendelian randomization analysis. Nat Genet 2020;52(11):1239–46.
10. Porcu E, Rueger S, Lepik K, et al. Mendelian randomization integrating GWAS and eQTL data reveals genetic determinants of complex and clinical traits. Nat Commun 2019;10(1):3300.
11. Yang JJ, Jia Li L, Williams K, et al. An efficient genome-wide association test for multivariate phenotypes based on the fisher combination function. BMC Bioinformatics 2016;17(1):19.
12. Mo C, Ye Z, Ke H, et al. A new Mendelian randomization method to estimate causal effects of multivariable brain imaging exposures. Pacific Symposium on Biocomputing 2022;2021:73–84.
13. Feng H, Mancuso N, Gusev A, et al. Leveraging expression from multiple tissues using sparse canonical correlation analysis and aggregate tests improves the power of transcriptome-wide association studies. PLoS Genet 2021;17(4):e1008973.
14. Zuber V, Colijn JM, Klaver C, et al. Selecting likely causal risk factors from high-throughput experiments using multivariable Mendelian randomization. Nat Commun 2020;11(1):29.
15. Zhu Z, Zheng Z, Zhang F, et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. Nat Commun 2018;9(1):224.
16. Sanderson E, Smith GD, Windmeijer F, et al. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. Int J Epidemiol 2019;48(3):713–27.
17. Rees JMB, Wood AM, Burgess S. Extending the MR-epger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy. Stat Med 2017;36(29):4705–18.
18. Liu Y, Chen S, Li Z, et al. ACAT: a fast and powerful p value combination method for rare-variant analysis in sequencing studies. Am J Hum Genet 2019;104(3):410–21.
19. Liu Y, Xie J. Cauchy combination test: a powerful test with analytic p-value calculation under arbitrary dependency structures. J Am Stat Assoc 2020;115(529):393–402.
20. Wu C, Bradley J, Li Y, et al. A gene-level methylome-wide association analysis identifies novel Alzheimer’s disease genes. Bioinformatics 2021;37(14):1933–40.
21. Wilson DJ. The harmonic mean p-value for combining dependent tests. Proc Natl Acad Sci 2019;116(4):1195–200.
22. Yang J, Keoki Williams L, Bui A. Identifying pleiotropic genes in genome-wide association studies from related subjects using the linear mixed model and fisher combination function. BMC Bioinformatics 2017;18(1):376.
23. Yurko R, Roeder K, Devlin B, et al. H-MAGMA, inheriting a shaky statistical foundation, yields excess false positives. Ann Hum Genet. 2021;85:97–100.
24. Kang H, Peck L, Keele L. Inference for instrumental variables: a randomization inference approach. J R Stat Soc A Stat Soc 2018;181(4):1231–54.
25. Burgess S, Bowden J. Integrating summarized data from multiple genetic variants in Mendelian randomization: bias and coverage properties of inverse-variance weighted methods. arXiv:1512.04486 [stat]. 2015.
26. Bowden J, Fabiola Del Greco M, Minelli C, et al. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. Stat Med 2017;36(11):1783–802.
27. Bowden J, Fabiola Del Greco M, Minelli C, et al. Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption. *Int J Epidemiol* 2019;48(3):728–42.

28. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol* 2016;40(7):597–608.

29. Pang Z, Chong J, Zhou G, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res* 2021;49(W1):W388–96.

30. Schwartzentruber J, Cooper S, Liu JZ, et al. Genome-wide meta-analysis, fine-mapping and integrative prioritization implicate new Alzheimer’s disease risk genes. *Nat Genet* 2021;53(3):392–402.

31. Kunkle BW, Grenier-Boley B, Sims R, et al. Genetic meta-analysis of diagnosed Alzheimer’s disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nat Genet* 2019;51(3):414–30.

32. Greenwood AK, Gockley J, Daily K, et al. Agora: an open platform for exploration of alzheimer’s disease evidence: genetics/omics and systems biology. *Alzheimers Dement* 2020;16:e046129.

33. ABCA7 - Alzforum. (n.d.). Retrieved 9 November 2021 https://www.alzforum.org/alzpedia/abca7.

34. Hollingworth P, Harold D, Sims R, et al. Common variants in ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer’s disease. *Nat Genet* 2011;43(5):429.

35. Naj AC, Jun G, Beecham GW, et al. Common variants in MS4A4/M54A6E, CD2uAP, CD33, and EPHA1 are associated with late-onset Alzheimer’s disease. *Nat Genet* 2011;43(5):436.

36. Carrasquillo MM, Crook JE, Pedraza O, et al. Late-onset Alzheimer risk variants in memory decline, incident mild cognitive impairment and Alzheimer disease. *Neurobiol Aging* 2015;36(1):60.

37. Ramirez LM, Goukasian N, Porat S, et al. Common variants in ABCA7 and MS4A6A are associated with cortical and hippocampal atrophy. *Neurobiol Aging* 2016;39:82–9.

38. Satoh K, Abe-Dohmae S, Yokoyama S, et al. ATP-binding cassette transporter A7 (ABCA7) loss of function alters Alzheimer amyloid processing. *J Biol Chem* 2015;290(40):24152.

39. Sakae N, Liu C-C, Shinohara M, et al. ABCA7 deficiency accelerates amyloid- generation and Alzheimer’s neuronal pathology. *J Neurosci* 2016;36(13):3848.

40. ATPIB1 - Sodium/potassium-transporting ATPase subunit beta-1 - Homo sapiens (Human) - ATP1B1 gene & protein. (n.d.). Retrieved 9 November 2021 https://www.uniprot.org/uniprot/P05026.

41. Bab-Dinitz E, Albeck S, Peleg Y, et al. C-terminal lobe of the β subunit of Na,K-ATPase and H,K-ATPase resembles cell adhesion molecules. *Biochemistry* 2009;48(36):8684–91.

42. Relton CL, Smith GD. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *Int J Epidemiol* 2012;41(1):161–76.

43. Zhao Q, Chen Y, Wang J, et al. Powerful three-sample genome-wide design and robust statistical inference in summary-data Mendelian randomization. *Int J Epidemiol* 2019;48(5):1478–92.

44. Huang Y-T, VanderWeele TJ, Lin X. Joint analysis of SNP and gene expression data in genetic association studies of complex diseases. *Ann Appl Stat* 2014;8(1):352–76.

45. Huang Y-T, Liang L, Moffatt MF, et al. iGWAS: integrative genome-wide association studies of genetic and genomic data for disease susceptibility using mediation analysis. *Genet Epidemiol* 2019;43(5):347–56.

46. Burgess S, Foley CN, Zuber V. Inferring causal relationships between risk factors and outcomes from genome-wide association study data. *Annu Rev Genomics Hum Genet* 2018;19(1):303–27.

47. Burgess S, Timpson NJ, Ebrahim S, et al. Mendelian randomization: where are we now and where are we going? *Int J Epidemiol* 2015;44(2):379–88.

48. Sieberts SK, Perumal TM, Carrasquillo MM, et al. Large eQTL meta-analysis reveals differing patterns between cerebral cortical and cerebellar brain regions. *Scientific Data* 2020;7(1):340.

49. Robins C, Liu Y, Fan W, et al. Genetic control of the human brain proteome. *Am J Hum Genet* 2021;108(3):400–10.