Transcriptome-wide association studies accounting for co-localization using Egger regression

Richard Barfield[1], Helian Feng[2], Alexander Gusev[3,4], Lang Wu[5], Wei Zheng[5], Bogdan Pasaniuc[6,7,8], Peter Kraft[2,9,10]

Author Affiliations

1) Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.
2) Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA
3) Division of Population Sciences, Dana-Farber Cancer Institute
4) Division of Genetics, Brigham & Women's Hospital
5) Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN, USA.
6) Department of Pathology & Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA
7) Bioinformatics Interdepartmental Program, University of California, Los Angeles, Los Angeles, CA, USA
8) Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA
9) Program in Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, Boston, MA, USA
10) Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

Corresponding Author: Peter Kraft (617-432-4271).

655 Huntington Avenue

Building II Room 249A

Boston, Massachusetts 02115

Abstract

Integrating genome-wide association (GWAS) and expression quantitative trait locus (eQTL) data into transcriptome-wide association studies (TWAS) based on predicted expression can boost power to detect novel disease loci or pinpoint the susceptibility gene at a known disease locus. However, it is often the case that multiple eQTL genes co-localize at disease loci, making the identification of the true susceptibility gene challenging, due to confounding through linkage disequilibrium (LD). To distinguish between true susceptibility genes (where the genetic effect on phenotype is mediated through expression) and co-localization due to LD, we examine an extension of the Mendelian Randomization Egger regression method that allows for LD while only requiring summary association data for both GWAS and eQTL. We derive the standard TWAS approach in the context of Mendelian Randomization and show in simulations that the standard TWAS does not control Type I error for causal gene identification when eQTLs have pleiotropic or LD-confounded effects on disease. In contrast, LD Aware MR-Egger regression can control Type I error in this case while attaining similar power of LD as other methods in situations where these provide valid tests. However, when the direct effects of genetic variants on traits are correlated with the eQTL associations, all of the methods we examined including LD Aware MR-Egger regression can have inflated Type I error. We illustrate these methods by integrating gene expression within a recent large-scale breast cancer GWAS to provide guidance on susceptibility gene identification.

Key Words: Genome-wide association study; Gene Expression, Mendelian Randomization, transciptome-wide association study

Introduction

Integrating data from genome-wide association studies (GWAS) of disease and expression quantitative trait loci (eQTL) studies can help detect novel disease loci and pinpoint genes of interest. Traditional studies have either matched the eQTL and GWAS association signals using ad-hoc overlap statistics or estimated the probability that the association signals are due to the same genetic variants (i.e., co-localization) [Giambartolomei, et al. 2014; Hormozdiari, et al. 2016; Wallace 2013]. More recent methods have proposed to test the association between transcript expression levels and disease risk by first using eQTL reference data to build multi-marker predictors of expression, and then testing the association between the genetically predicted expression and disease risk in a large GWAS [Gamazon, et al. 2015; Gusev, et al. 2016]. The latter methods, which we jointly refer to as Transcriptome-Wide Association Studies (TWAS), have been extended to the case when only summary statistics are available from the disease GWAS, the eQTL study, or both [Barbeira, et al. 2016; Gusev, et al. 2016; Zhu, et al. 2016].

Although TWAS has been successful in identifying many genes whose genetically regulated expression is associated to traits, a major limitation of TWAS is that it cannot distinguish between a causal effect of expression on disease and a tagging association within the same region due to correlations among SNPs (i.e. linkage disequilibrium, LD) [Mancuso, et al. 2017a; Wainberg, et al. 2017]. A SNP used in expression prediction of gene A may be in linkage disequilibrium (LD) with nearby SNPs used in the prediction of gene B for that is not involved in disease etiology. This LD will induce association between the genetically regulated expression of gene B and disease, causing the TWAS test statistic to reject the null of no association between predicted expression and disease at gene B even though expression levels are not causally related to disease; this effect is similar to standard LD-tagging in GWAS where LD induces significant association statistics at non-causal SNPs [Mancuso, et al. 2017a]. As TWAS methods were originally proposed as tests for association between local genetically regulated component of expression and disease with no causality guarantees [Gamazon, et al. 2015; Gusev, et al. 2016; Mancuso, et al. 2017a; Mancuso, et al. 2017b; Zhu, et al. 2016], it remains unclear whether and when TWAS can be interpreted as valid tests of causality. In contrast to TWAS, co-localization

analyses, including COLOC, eCAVIAR, focus on estimating the probability of the SNPs causals for eQTL to be the same as for GWAS, irrespective of direction of genetic effect on expression or disease, and are not designed to test for an effect of gene expression on disease risk [Giambartolomei, et al. 2014; Hormozdiari, et al. 2016; Wallace 2013].

In this work, we explore the use of TWAS methods in the context of causal gene identification. As opposed to the standard TWAS that mainly focused on association testing to identify new genomic regions harboring disease genes, in this work we investigate the utility of TWAS methods in the context of causal gene localization (which utilizes a much more stringent definition of true/false positive). We re-derive TWAS methods in a Mendelian Randomization (MR) framework and show that the standard TWAS statistic is a special case of MR that use eQTLs as genetic instruments to test the causal association between expression and disease. We leverage the growing literature on methods for MR that use summary statistics for both the associations between single nucleotide polymorphisms (SNPs) and the intermediate trait and between SNPs and disease; that use multi-SNP genetic instruments [Burgess, et al. 2016]; and that remain valid when some of the assumptions underlying standard MR are violated [Bowden, et al. 2015]. In particular, the MR Egger regression approach relaxes the assumption that the association between genetic instruments and disease is only mediated through the intermediate trait—which would not be the case if the SNPs in the genetic instrument had pleiotropic effects, for example [Bowden, et al. 2015]. We investigate the use of a variant of MR Egger regression that accounts for LD among the variants used in the genetic predictor for gene expression (LD aware MR Egger regression) in the context of TWAS. This method was first proposed independently of this group in the discussion section of Burgess and Thompson 2017 with derivations provided in their appendix [Burgess and Thompson 2017]. We use extensive simulations to compare the performance of LD aware MR Egger to other TWAS and MR statistics and show that it remains valid under specific violations of the assumptions underlying MR, while retaining comparable power to the other approaches when the assumptions do hold.

The structure of this paper is as follows. First, we introduce the conceptual model relating the outcome and gene expression to the SNPs. Second, we discuss four existing approaches for testing the association between the outcome and gene expression. We show that the traditional TWAS test statistic of Gusev et al. [Gusev, et al. 2016] is equivalent to an LD aware version of

standard MR using summary statistics and compare this approach to LD aware MR-Egger regression (LDA MR-Egger) [Burgess and Thompson 2017]. Third, we examine the statistical properties of these estimates and their performance via simulation in the presence or absence of a direct effect of SNPs on disease. To the best of our knowledge, this is the first time the empirical performance of LDA MR-Egger regression has been examined in this context. We apply the various approaches to summary statistics from a GWAS on Breast Cancer [Michailidou, et al. 2017] with eQTL data from a breast tissue panel in GTEx [Consortium 2013]. We conclude by providing guidance on the interpretability of standard TWAS and MR tests can be interpreted as valid tests in the context of eQTL and GWAS integration.

Methods

Main Models

Let $\boldsymbol{Y}$ denote the outcome (*nx1*), $\boldsymbol{M}$ the mediator (*nx1*), and $\boldsymbol{G}$ the SNP matrix (*n*x*J*) of interest. We assume that the columns of $\boldsymbol{G}$ have been standardized to have mean zero and variance one. In the TWAS setting, $\boldsymbol{M}$ is gene expression. We denote the LD structure of the SNPs $\boldsymbol{G}$ as $\boldsymbol{\Sigma}$, a *JxJ* symmetric positive definite matrix. (In practice, this matrix may be a modified version of the empirical LD matrix, adding a small ridge regularization to ensure the matrix is symmetric positive definite and invertible [Pasaniuc and Price 2017].) For a link function *g*, we have the following conceptual models relating $\boldsymbol{M}$, $\boldsymbol{G}$, and $\boldsymbol{Y}$.

$$g\big(E(\boldsymbol{Y}|\boldsymbol{M},\boldsymbol{G})\big) = \gamma_0 + \boldsymbol{M}\gamma + \boldsymbol{G}\boldsymbol{\theta} \qquad \textbf{(i.a.)}$$

$$g\big(E(\boldsymbol{Y}|\boldsymbol{G})\big) = \gamma_0^* + \boldsymbol{G}\boldsymbol{\beta}_G \qquad \textbf{(i.b.)}$$

$$\boldsymbol{M} = \beta_0 + \boldsymbol{G}\boldsymbol{\beta}_E + \boldsymbol{\epsilon}_M; \boldsymbol{\epsilon}_M \sim N(0, \boldsymbol{I}_n\sigma^2) \qquad \textbf{(i.c.)}$$

In the above model, $\boldsymbol{\theta}$ is a *J*-column vector of $\boldsymbol{G}$ effects on $\boldsymbol{Y}$ conditional on $\boldsymbol{M}$, γ is the effect of $\boldsymbol{M}$ on $\boldsymbol{Y}$ conditional on $\boldsymbol{G}$, $\boldsymbol{\beta}_E$ is the *J*-column vector of SNP effects on the mediator, and $\boldsymbol{\beta}_G$ is the *J*-vector of the $\boldsymbol{G}$ effects marginal over $\boldsymbol{M}$. $\boldsymbol{\epsilon}_M$ and $\boldsymbol{I}_n$ represent the residual variance in $\boldsymbol{M}$ and the *n* x *n* identity matrix respectively. We are interested in the situation where we cannot directly estimate the parameters in model (*i.a*), as we do not have complete data on $\boldsymbol{Y}$, $\boldsymbol{M}$ and $\boldsymbol{G}$

from (sufficiently many) individuals. We want to place inference on γ, because if $\gamma \neq 0$, gene expression $\boldsymbol{M}$ affects the trait. To relate the parameters in the marginal model of $\boldsymbol{Y}$ and the marginal model of $\boldsymbol{M}$ we assume one of the following for the remainder of the paper:

1) g is either the log or identity link function.
2) **Y** is a sufficiently rare binary trait and g is the logit link.

If either of the two conditions above hold, we will have the following:

$$\boldsymbol{\beta}_G \approx \boldsymbol{\beta}_E \gamma + \boldsymbol{\theta}. \qquad \textbf{(ii.)}$$

If g is the log or linear link, the approximation will be exact. Equation (ii) suggests that the effects of $\boldsymbol{G}$ marginal over $\boldsymbol{M}$ are a function of the eQTL parameters ($\boldsymbol{\beta}_E$), the effect of the gene expression on the outcome conditional on $\boldsymbol{G}$ (γ), and the effect of $\boldsymbol{G}$ on the outcome conditional on M ($\boldsymbol{\theta}$). We will call $\boldsymbol{\theta}$ the direct effect and $\boldsymbol{\beta}_E \gamma$ the mediated effect. If the necessary causal assumptions are met, these parameters could be interpreted in the causal/counterfactual framework [Valeri and VanderWeele 2013], but for this work it is not necessary.

In practice, we do not have an overlap of data on $\boldsymbol{M}$, $\boldsymbol{Y}$, and $\boldsymbol{G}$. Instead, we have a sample of size N that the GWAS is run on to estimate $\boldsymbol{\beta}_G$ and an independent sample of size N_E that was used to estimate $\boldsymbol{\beta}_E$. We therefore cannot estimate $\boldsymbol{\theta}$ directly. Moreover, we typically only have estimates of individual SNP effects marginal over the other SNPs, $\widehat{\boldsymbol{\beta}}_E^*$ and $\widehat{\boldsymbol{\beta}}_G^*$. For our purposes, $\hat{\beta}_{G,j}^*$ and $\hat{\beta}_{E,j}^*$ were estimated with the same reference allele for SNP j. If they were not, the sign of the effect can be changed so as to refer to the same reference allele.

The formulas given above relating the mean of $\boldsymbol{Y}$ and $\boldsymbol{M}$ to $\boldsymbol{G}$ were given on the conditional level. We therefore transform our marginal estimates ($\widehat{\boldsymbol{\beta}}_E^*$ and $\widehat{\boldsymbol{\beta}}_G^*$) to the conditional scale. Given an estimate of the LD matrix ($\boldsymbol{\Sigma}$) we estimate the conditional eQTL and GWAS effects as $\widehat{\boldsymbol{\beta}}_E = \boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\beta}}_E^*$ and $\widehat{\boldsymbol{\beta}}_G = \boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\beta}}_G^*$ [Shi, et al. 2016]. If the marginal effect estimates were not calculated on the standardized genotypes they can be transformed by multiplying by the square root of $2p_j(1-p_j)$, where p_j is the MAF of SNP j. We assume that $\widehat{\boldsymbol{\beta}}_E$ and $\widehat{\boldsymbol{\beta}}_G$ are unbiased for $\boldsymbol{\beta}_E$ and $\boldsymbol{\beta}_G$ respectively.

We do not know the form of $\boldsymbol{\theta}$; it could be a constant or vary by SNP. All we know is that it is a vector of length J. Our estimated GWAS effects, (given our assumptions above) are a

function of $\boldsymbol{\theta}, \boldsymbol{\beta}_E, \gamma$ and the sampling error:

$$\widehat{\boldsymbol{\beta}}_G \approx \boldsymbol{\theta} + \boldsymbol{\beta}_E\gamma + \boldsymbol{\epsilon_G}, \boldsymbol{\epsilon_G} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_G).$$

If the SNPs are not in LD, then the marginal and the conditional will be equal ($\widehat{\boldsymbol{\beta}}_G \approx \widehat{\boldsymbol{\beta}}_G^*$).

We next derive the covariance of $\widehat{\boldsymbol{\beta}}_G$ ($\boldsymbol{\Sigma}_G$). Let $\boldsymbol{\sigma}_G^*$ denote a *Jx1* vector of the marginal standard errors of $\widehat{\boldsymbol{\beta}}_G^*$. Let "." denote element wise multiplication between two matrices. As $\boldsymbol{G}$ has been standardized, $cov(\widehat{\boldsymbol{\beta}}_G^*) = \boldsymbol{\Sigma} \cdot \boldsymbol{\sigma}_G^* \boldsymbol{\sigma}_G^{*\boldsymbol{T}}$. This gives the covariance of our conditional GWAS estimates:

$$\boldsymbol{\Sigma}_G = cov(\widehat{\boldsymbol{\beta}}_G) = cov(\boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\beta}}_G^*) = \boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma} \cdot \boldsymbol{\sigma}_G^* \boldsymbol{\sigma}^{*T}_G)\boldsymbol{\Sigma}^{-1}.$$

If $\boldsymbol{\sigma}_G^* = v\mathbf{1}_J$, $\mathbf{1}_J$ is a column vector of ones, then $\boldsymbol{\Sigma}_G = \boldsymbol{\Sigma}^{-1}v^2$. The main goal is to derive a valid test of the null hypothesis γ=0 and potentially estimate γin the presence of direct effects of the SNPs onto the outcome. We next go over some common methods for testing for an association between gene expression and outcome.

Transcriptome Wide Association Studies (TWAS)

The standard TWAS statistic uses summary statistics to test for an association between genetically predicted gene expression and a phenotype of interest [Gusev, et al. 2016]. The TWAS does not necessarily estimate the γ above but does provide a valid test for the association between the gene of interest and the outcome. The TWAS test statistic is:

$$Z_{TWAS} = \frac{\widehat{\boldsymbol{\beta}}_E^{*\boldsymbol{T}} \boldsymbol{\Sigma}^{-1} \boldsymbol{Z}_G^*}{\sqrt{\widehat{\boldsymbol{\beta}}_E^{*\boldsymbol{T}} \boldsymbol{\Sigma}^{-1} \widehat{\boldsymbol{\beta}}_E^*}},$$

where $\boldsymbol{Z}_G^* = \widehat{\boldsymbol{\beta}}_G^* / \boldsymbol{\sigma}_G^*$, is a column vector of the marginal test statistics. The test statistic is then compared to a standard normal to assess significance. The TWAS can use either the marginal or conditional eQTL estimates as weights. If using the conditional, the equation above becomes:

$$Z_{TWAS} = \frac{\widehat{\boldsymbol{\beta}}_E^{\boldsymbol{T}} \boldsymbol{Z}_G^*}{\sqrt{\widehat{\boldsymbol{\beta}}_E^{\boldsymbol{T}} \boldsymbol{\Sigma} \widehat{\boldsymbol{\beta}}_E}}.$$

The Summary Mendelian Randomization Estimator

The summary MR estimator [Johnson 2011] is:

$$\hat{\gamma}_{MR} = \frac{\widehat{\boldsymbol{\beta}}_E^{*\boldsymbol{T}} \boldsymbol{V}^{-1} \widehat{\boldsymbol{\beta}}_G^*}{\widehat{\boldsymbol{\beta}}_E^{*\boldsymbol{T}} \boldsymbol{V}^{-1} \widehat{\boldsymbol{\beta}}_E^*},$$

where $\boldsymbol{V}$ is a diagonal matrix with $v_{jj} = \sigma_{G,j}^{2*} = var(\hat{\beta}_{G,j}^*)$. If there are no direct effects, i.e. $\boldsymbol{\theta} = 0$, and the SNPs are not in LD with each other ($\boldsymbol{\Sigma} = \boldsymbol{I}_J$), then this will be an unbiased estimate of γ. The MR estimate is best suited for analysis where the mediator is another phenotype and the SNPs are taken from across the genome. This estimate (which we term as the MR estimate) can be rewritten as:

$$\hat{\gamma}_{MR} = \frac{\sum_{j=1}^{J} \hat{\beta}_{E,j}^* \hat{\beta}_{G,j}^* \sigma_{G,j}^{-2*}}{\sum_{j=1}^{J} \hat{\beta}_{E,j}^{2*} \sigma_{G,j}^{-2*}}.$$

We estimate its variance as:

$$var(\hat{\gamma}_{MR}) = \frac{\hat{\sigma}_{MR}^2}{\sum_{j=1}^{J} \hat{\beta}_{E,j}^{2*} \sigma_{G,j}^{-2*}},$$

$$\hat{\sigma}_{MR}^2 = \frac{1}{J-1} \sum_{j=1}^{J} \sigma_{G,j}^{-2*} \left(\hat{\beta}_{G,j}^* - \hat{\beta}_{E,j}^* \hat{\gamma}_{MR}\right)^2.$$

The MR test statistic is then:

$$Z_{MR} = \frac{\hat{\gamma}_{MR}}{\sqrt{var(\hat{\gamma}_{MR})}} = \frac{\sum_{j=1}^{J} \hat{\beta}_{E,j}^* \hat{\beta}_{G,j}^* \sigma_{G,j}^{-2*}}{\hat{\sigma}_{MR} \sqrt{\sum_{j=1}^{J} \hat{\beta}_{E,j}^{2*} \sigma_{G,j}^{-2*}}}.$$

For testing, we compare Z_{MR} to the quantiles of a *t*-distribution with *J*-1 degrees of freedom. If the SNPs are in LD or if there are direct effects, $\hat{\gamma}_{MR}$ can lead to incorrect inference. The MR estimate is a weighted linear regression without an intercept of the marginal GWAS estimates on the marginal eQTL estimates with weights equal to $\sigma_{G,j}^{-2*}$.

We note here that Zhu et al. refer to their single SNP MR test as a summary Mendelian Randomization, that approach is not the same as what is presented here [Zhu, et al. 2016]. The approach above takes into account all SNPs at once as opposed to doing one SNP at a time. The

Zhu et al approach provides an estimate for pleiotropy (HEIDI) based on the summary statistics and conditioning on the lead SNP.

MR-Egger Estimate

If the MR estimate is a weighted linear regression without an intercept, the MR-Egger extends the MR by including an intercept (α) to the weighted linear regression. It assumes that $E\left(\widehat{\boldsymbol{\beta}}_G^* \middle| \widehat{\boldsymbol{\beta}}_E^*\right) = \alpha \mathbf{1}_J + \widehat{\boldsymbol{\beta}}_E^* \gamma$. The estimates are (same $\boldsymbol{V}$ as the MR estimate):

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\gamma}_{MRE} \end{bmatrix} = \begin{bmatrix} \mathbf{1}_J^T \boldsymbol{V}^{-1} 1_J & \mathbf{1}_J^T \boldsymbol{V}^{-1} \widehat{\boldsymbol{\beta}}_E^* \\ \widehat{\boldsymbol{\beta}}_E^{*T} \boldsymbol{V}^{-1} 1_J & \widehat{\boldsymbol{\beta}}_E^{*T} \boldsymbol{V}^{-1} \widehat{\boldsymbol{\beta}}_E^* \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{1}_J^T \boldsymbol{V}^{-1} \\ \widehat{\boldsymbol{\beta}}_E^{*T} \boldsymbol{V}^{-1} \end{bmatrix} \widehat{\boldsymbol{\beta}}_G^*,$$

The estimate of $\gamma$is then:

$$\hat{\gamma}_{MRE} = \frac{\left(\sum_{j=1}^{J} \sigma_{G,j}^{-2*}\right)\left(\sum_{j=1}^{J} \hat{\beta}_{E,j}^* \hat{\beta}_{G,j}^* \sigma_{G,j}^{-2*}\right) - \left(\sum_{j=1}^{J} \hat{\beta}_{E,j} \sigma_{G,j}^{-2*}\right)\left(\sum_{j=1}^{J} \hat{\beta}_{G,j}^* \sigma_{G,j}^{-2*}\right)}{\left(\sum_{j=1}^{J} \sigma_{G,j}^{-2*}\right)\left(\sum_{j=1}^{J} \hat{\beta}_{E,j}^{*2} \sigma_{G,j}^{-2*}\right) - \left(\sum_{j=1}^{J} \hat{\beta}_{E,j}^* \sigma_{G,j}^{-2*}\right)^2}.$$

The variance of the estimate is:

$$var(\hat{\gamma}_{MRE}) = \hat{\sigma}_{MRE}^2 \frac{\sum_{j=1}^{J} \sigma_{G,j}^{-2*}}{\left(\sum_{j=1}^{J} \sigma_{G,j}^{-2*}\right)\left(\sum_{j=1}^{J} \hat{\beta}_{E,j}^{*2} \sigma_{G,j}^{-2*}\right) - \left(\sum_{j=1}^{J} \hat{\beta}_{E,j}^* \sigma_{G,j}^{-2*}\right)^2},$$

$$\hat{\sigma}_{MRE}^2 = \frac{1}{J-2} \sum_{j=1}^{J} \sigma_{G,j}^{-2*} \left(\hat{\beta}_{G,j}^* - \hat{\alpha} - \hat{\beta}_{E,j}^* \hat{\gamma}_{MRE}\right)^2.$$

The MR-Egger test statistic is then:

$$Z_{MRE} = \frac{\left(\sum_{j=1}^{J} \sigma_{G,j}^{-2*}\right)\left(\sum_{j=1}^{J} \hat{\beta}_{E,j}^* \hat{\beta}_{G,j}^* \sigma_{G,j}^{-2*}\right) - \left(\sum_{j=1}^{J} \hat{\beta}_{E,j} \sigma_{G,j}^{-2*}\right)\left(\sum_{j=1}^{J} \hat{\beta}_{G,j}^* \sigma_{G,j}^{-2*}\right)}{\hat{\sigma}_{MRE} \sqrt{\sum_{j=1}^{J} \sigma_{G,j}^{-2*}} \sqrt{\left(\sum_{j=1}^{J} \sigma_{G,j}^{-2*}\right)\left(\sum_{j=1}^{J} \hat{\beta}_{E,j}^{*2} \sigma_{G,j}^{-2*}\right) - \left(\sum_{j=1}^{J} \hat{\beta}_{E,j}^* \sigma_{G,j}^{-2*}\right)^2}}.$$

To test, we compare to the quantiles of a *t*-distribution with *J*-2 degrees of freedom. If the SNPs are in LD, the test for $\hat{\gamma}_{MRE}$ may lead to incorrect inference due to the variance of $\hat{\gamma}_{MRE}$ being misspecified.

LD-Aware MR (LDA MR) Estimate

The LDA MR estimator of γ extends the MR estimator by relaxing the assumption of the SNPs being independent. It however, still requires that there are no direct effects. Recall that $\boldsymbol{\Sigma}_G = cov(\widehat{\boldsymbol{\beta}}_G)$. The LDA MR estimator is then:

$$\hat{\gamma}_{LDMR} = \frac{\widehat{\boldsymbol{\beta}}_E^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_G}{\widehat{\boldsymbol{\beta}}_E^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_E},$$

And we estimate the variance as:

$$var(\hat{\gamma}_{LDMR}) = \frac{\hat{\sigma}_{LDMR}^2}{\widehat{\boldsymbol{\beta}}_E^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_E},$$

$$\hat{\sigma}_{LDMR}^2 = \frac{1}{J-1}\left(\widehat{\boldsymbol{\beta}}_G - \widehat{\boldsymbol{\beta}}_E \hat{\gamma}_{LDMR}\right)^T \boldsymbol{\Sigma}_G^{-1}\left(\widehat{\boldsymbol{\beta}}_G - \widehat{\boldsymbol{\beta}}_E \hat{\gamma}_{LDMR}\right).$$

The test statistic is then:

$$Z_{LDMR} = \frac{\hat{\gamma}_{LDA-MR}}{\sqrt{var(\hat{\gamma}_{LDMR})}} = \frac{\widehat{\boldsymbol{\beta}}_E^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_G}{\hat{\sigma}_{LDMR}\sqrt{\widehat{\boldsymbol{\beta}}_E^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_E}}.$$

Similar to the MR, we compare to a *t* distribution with *J*-1 df. If there are direct effects, $\hat{\gamma}_{LDMR}$ can lead to incorrect inference. Just as the MR estimate is a weighted linear regression without an intercept, the LDA MR estimate is a weighted linear regression without an intercept using the weight matrix $\boldsymbol{\Sigma}_G^{-1}$. If $\sigma_{G,j}^* = v$ for all SNPs, the LDA MR and the TWAS test-statistic run on the marginal eQTL will be proportional to each other by $\hat{\sigma}_{LDMR}^{-1}$ (Appendix A).

LDA MR-Egger Estimate

LDA MR-Egger regression extends the MR-Egger approach to incorporate the LD structure of the SNPs. We include an intercept, in the aims of accounting for the direct effect of the SNPs. The estimates of the intercept and γ are:

$$\begin{bmatrix} \hat{\alpha}_{LD} \\ \hat{\gamma}_{LDMRE} \end{bmatrix} = \left(\begin{bmatrix} \mathbf{1}_J^T \boldsymbol{\Sigma}_G^{-1} \mathbf{1}_J & \mathbf{1}_J^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_E \\ \widehat{\boldsymbol{\beta}}_E^T \boldsymbol{\Sigma}_G^{-1} \mathbf{1}_J & \widehat{\boldsymbol{\beta}}_E^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_E \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{1}_J^T \boldsymbol{\Sigma}_G^{-1} \\ \widehat{\boldsymbol{\beta}}_E^T \boldsymbol{\Sigma}_G^{-1} \end{bmatrix} \widehat{\boldsymbol{\beta}}_G,$$

with γ more succinctly being written as:

$$\hat{\gamma}_{LDMRE} = \frac{\left(\mathbf{1}_J \boldsymbol{\Sigma}_G^{-1} \mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_G\right) - \left(\mathbf{1}_J^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_E\right)\left(\mathbf{1}_J^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_G\right)}{\left(\mathbf{1}_J^T \boldsymbol{\Sigma}_G^{-1} \mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_E\right) - \left(\mathbf{1}_J^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_E\right)^2}.$$

The variance is estimated as:

$$var(\hat{\gamma}_{LDMRE}) = \hat{\sigma}_{LDMRE}^2 \frac{\mathbf{1}_J \boldsymbol{\Sigma}_G^{-1} \mathbf{1}_J}{\left(\mathbf{1}_J^T \boldsymbol{\Sigma}_G^{-1} \mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_E\right) - \left(\mathbf{1}_J^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_E\right)^2},$$

$$\hat{\sigma}_{LDMRE}^2 = \frac{1}{J-2}\left(\widehat{\boldsymbol{\beta}}_G - \hat{\alpha}_{LD}\mathbf{1}_J - \widehat{\boldsymbol{\beta}}_E \hat{\gamma}_{LDMRE}\right)^T \boldsymbol{\Sigma}_G^{-1}\left(\widehat{\boldsymbol{\beta}}_G - \hat{\alpha}_{LD}\mathbf{1}_J - \widehat{\boldsymbol{\beta}}_E \hat{\gamma}_{LDMRE}\right).$$

The test statistic is then:

$$Z_{LDMRE} = \frac{\hat{\gamma}_{LDMRE}}{\sqrt{var(\hat{\gamma}_{LDMRE})}} = \frac{\left(\mathbf{1}_J \boldsymbol{\Sigma}_G^{-1} \mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_G\right) - \left(\mathbf{1}_J^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_E\right)\left(\mathbf{1}_J^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_G\right)}{\hat{\sigma}_{LDMRE}\sqrt{\mathbf{1}_J \boldsymbol{\Sigma}_G^{-1} \mathbf{1}_J}\sqrt{\left(\mathbf{1}_J^T \boldsymbol{\Sigma}_G^{-1} \mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_E\right) - \left(\mathbf{1}_J^T \boldsymbol{\Sigma}_G^{-1} \widehat{\boldsymbol{\beta}}_E\right)^2}}.$$

For testing, we compare to a *t*-distribution with *J*-2 degrees of freedom. This estimate is a combination of the approaches from LDA MR and the MR-Egger. The LDA MR-Egger will provide valid inference in the same scenarios as the MR-Egger, but in addition when the SNPs are in LD. Details of all tests are provided in Table I.

Biases of Estimates

We first focus on the estimates that do not incorporate an intercept, the MR and the LDA MR. We now examine the bias of the MR estimate:

$$\hat{\gamma}_{MR} = \frac{\widehat{\boldsymbol{\beta}}_E^{*T} \boldsymbol{V}^{-1} \widehat{\boldsymbol{\beta}}_G^*}{\widehat{\boldsymbol{\beta}}_E^{*T} \boldsymbol{V}^{-1} \widehat{\boldsymbol{\beta}}_E^*},$$

Note that $\widehat{\boldsymbol{\beta}}_G^* = \boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_G$ and $\widehat{\boldsymbol{\beta}}_E^* = \boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E$. Thus $E(\widehat{\boldsymbol{\beta}}_G^*)$=E($\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_G$)=$\boldsymbol{\Sigma}\boldsymbol{\beta}_E\gamma + \boldsymbol{\Sigma}\boldsymbol{\theta}$. We assume that $\widehat{\boldsymbol{\beta}}_E$ and $\widehat{\boldsymbol{\beta}}_G$ were estimated from different samples and thus are independent, E($\widehat{\boldsymbol{\beta}}_G|\widehat{\boldsymbol{\beta}}_E$)=E($\widehat{\boldsymbol{\beta}}_G$). Using this gives us that:

$$E\left(\hat{\gamma}_{MR}\middle|\widehat{\boldsymbol{\beta}}_E^*\right) = \gamma\frac{\widehat{\boldsymbol{\beta}}_E^{*T} \boldsymbol{V}^{-1} \boldsymbol{\Sigma}\boldsymbol{\beta}_E}{\widehat{\boldsymbol{\beta}}_E^{*T} \boldsymbol{V}^{-1} \widehat{\boldsymbol{\beta}}_E^*} + \frac{\widehat{\boldsymbol{\beta}}_E^{*T} \boldsymbol{V}^{-1} \boldsymbol{\Sigma}\boldsymbol{\theta}}{\widehat{\boldsymbol{\beta}}_E^{*T} \boldsymbol{V}^{-1} \widehat{\boldsymbol{\beta}}_E^*}.$$

Then use that $\widehat{\boldsymbol{\beta}}_E^* = \boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E$:

$$E\left(\hat{\gamma}_{MR}\middle|\widehat{\boldsymbol{\beta}}_E^*\right) = \gamma\frac{\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\boldsymbol{\beta}_E}{\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E} + \frac{\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\boldsymbol{\theta}}{\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E}$$

The term $\widehat{\boldsymbol{\beta}}_E$ was estimated from a sample size of N_E. As $N_E \to\infty$, we again have that $\widehat{\boldsymbol{\beta}}_E \to \boldsymbol{\beta}_E$ and the first term goes to γ. Assuming N_E is sufficiently large:

$$E\left(\hat{\gamma}_{MR}\middle|\widehat{\boldsymbol{\beta}}_E^*\right) \approx \gamma + \frac{\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\boldsymbol{\theta}}{\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E}.$$

Now unless $\boldsymbol{\theta} = 0$ or the transformation $\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}$ is orthogonal to $\boldsymbol{\theta}$, the MR will be biased. For the LDA MR:

$$\hat{\gamma}_{LDMR} = \frac{\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_G}{\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E},$$

$$E\left(\hat{\gamma}_{LDMR}\middle|\widehat{\boldsymbol{\beta}}_E\right) = \gamma\frac{\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\boldsymbol{\beta}_E}{\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E} + \frac{\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\boldsymbol{\theta}}{\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E}.$$

As $N_E \to\infty$, $\widehat{\boldsymbol{\beta}}_E \to \boldsymbol{\beta}_E$ and the first term will go to γ. Assuming that N_E is sufficiently large enough for this to occur, we have:

$$E\left(\hat{\gamma}_{LDMR}\middle|\hat{\beta}_E\right) \approx \gamma + \frac{\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\boldsymbol{\theta}}{\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E}.$$

Again, if $\boldsymbol{\theta} \neq 0$ or if the transformation $\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}$ is not orthogonal to $\boldsymbol{\theta}$, then the estimate will be biased. As J increases, the N_E needed for $\widehat{\boldsymbol{\beta}}_E \to \boldsymbol{\beta}_E$ will also increase.

We next assessed methods that include an intercept. For the MR-Egger estimate, using the fact that $\widehat{\boldsymbol{\beta}}_G^* = \boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_G$ and $\widehat{\boldsymbol{\beta}}_E^* = \widehat{\boldsymbol{\beta}}_E$, we have:

$$\hat{\gamma}_{MRE} = \frac{\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^{T*}\boldsymbol{V}^{-1}\widehat{\boldsymbol{\beta}}_G^*\right) - \left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\widehat{\boldsymbol{\beta}}_E^*\right)\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\widehat{\boldsymbol{\beta}}_G^*\right)}{\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^{T*}\boldsymbol{V}^{-1}\widehat{\boldsymbol{\beta}}_E^*\right) - \left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\widehat{\boldsymbol{\beta}}_E^*\right)^2}$$

$$= \frac{\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_G\right) - \left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E\right)\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_G\right)}{\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E\right) - \left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E\right)^2}.$$

$$E\left(\hat{\gamma}_{MRE}\middle|\widehat{\boldsymbol{\beta}}_E^*\right) = \gamma\frac{\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\boldsymbol{\beta}_E\right) - \left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E\right)\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\boldsymbol{\beta}_E\right)}{\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E\right) - \left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E\right)^2} +$$

$$\frac{\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\boldsymbol{\theta}\right) - \left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E\right)\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\boldsymbol{\theta}\right)}{\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E\right) - \left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E\right)^2}$$

As we have done previously, we assume N_E is sufficiently large so that $\widehat{\boldsymbol{\beta}}_E \approx \boldsymbol{\beta}_E$, making the first term γ.

$$E\left(\hat{\gamma}_{MRE}\middle|\widehat{\boldsymbol{\beta}}_E^*\right) \approx \gamma + \frac{\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\boldsymbol{\theta}\right) - \left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E\right)\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\boldsymbol{\theta}\right)}{\left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E\right) - \left(\mathbf{1}_J^T\boldsymbol{V}^{-1}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_E\right)^2}$$

Thus, the MR-Egger estimate will be unbiased if there is no direct effect ($\boldsymbol{\theta} = 0$), or if $\boldsymbol{\theta}$ is a constant, or if the numerator in the second term is equal to 0. Note that if $\boldsymbol{\Sigma}$=**I** and $\boldsymbol{V} \propto \frac{1}{N}\mathbf{I}$, as is the case in the setting where the MR-Egger estimate was originally proposed, then the numerator is equal to the empirical covariance between $\widehat{\boldsymbol{\beta}}_E$ and $\boldsymbol{\theta}$. This condition is referred to as the Instrument Strength Independent of Direct Effect (InSIDE) condition [Bowden, et al. 2015]. If the number of SNPs in the instrument for the mediator (J) is large, and there is no systematic relationship between $\boldsymbol{\beta}_E$ and $\boldsymbol{\theta}$, then this empirical covariance will be small. However, if J is small, as may be the case when gene expression is the mediator and the genetic instrument is limited to cis SNPs, then the numerator in the second term may be non-zero even when there is no systematic relationship between $\boldsymbol{\beta}_E$ and $\boldsymbol{\theta}$. Even if the INSIDE condition holds but $\boldsymbol{\Sigma} \neq \mathbf{I}$, the MR-Egger estimate will lead to improper inference, as it does not account for linkage disequilibrium.

We now consider the LDA MR-Egger estimate:

$$\hat{\gamma}_{LDMRE} = \frac{\left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_G\right) - \left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E\right)\left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_G\right)}{\left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E\right) - \left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E\right)^2},$$

Again, using that $\widehat{\boldsymbol{\beta}}_E$ and $\widehat{\boldsymbol{\beta}}_G$ were estimated from different studies, we have that

$$E\left(\hat{\gamma}_{LDMRE}\middle|\widehat{\boldsymbol{\beta}}_E\right) = \gamma\frac{\left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\mathbf{1}_J\right)\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\boldsymbol{\beta}_E - \left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E\right)\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\boldsymbol{\beta}_E}{\left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E\right) - \left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E\right)^2}$$

$$+\frac{\left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\mathbf{1}_J\right)\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\boldsymbol{\theta}-\left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E\right)\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\boldsymbol{\theta}}{\left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E\right)-\left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E\right)^2}.$$

As with the other estimates: $N_E \to\infty$, $\widehat{\boldsymbol{\beta}}_E \to \boldsymbol{\beta}_E$. If we assume that N_E is sufficiently large:

$$E\left(\hat{\gamma}_{LDMRE}\middle|\widehat{\boldsymbol{\beta}}_E\right) \approx \gamma+\frac{\left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\mathbf{1}_J\right)\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\boldsymbol{\theta}-\left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E\right)\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\boldsymbol{\theta}}{\left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\mathbf{1}_J\right)\left(\widehat{\boldsymbol{\beta}}_E^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E\right)-\left(\mathbf{1}_J^T\boldsymbol{\Sigma}_G^{-1}\widehat{\boldsymbol{\beta}}_E\right)^2}.$$

The numerator of the second term is a function of the sample univariate covariance between $\widehat{\boldsymbol{\beta}}_E$ and $\boldsymbol{\theta}$ weighted by $\boldsymbol{\Sigma}_G^{-1}$. The LDA MR-Egger estimate will be unbiased in the same situations as the MR-Egger estimate, but it does not require the SNPs be independent.

Simulation Study

We examine the empirical performance of four methods: the MR, MR-Egger, LDA MR, and the LDA MR-Egger. We note here that the MR and LDA MR approaches' empirical performance in the presence of LD has been examined in detail in Burgess et al. [2016], but for the sake of comparison we include them in our analysis [Burgess, et al. 2016]. Under our simulation scenario, the TWAS and the LDA MR test statistics are approximately the same ($\sigma^*_{G,j}$ is a constant), so we only present results on LDA MR. We generated the data such that all SNPs have a small causal effect. For each simulation, we generated summary eQTL and GWAS statistics from a multivariate normal distribution as opposed to individual level data [Han, et al. 2009]. We fixed the sample of the eQTL study to 1000 (N_E) and the sample size of the GWAS to 5000 (N). We varied the number of SNPs at the locus (J); the proportion of variation in $\boldsymbol{Y}$ explained by $\boldsymbol{G}$ ($h^2_{G\to Y}$) and $\boldsymbol{M}$ ($h^2_{E\to Y}$); the proportion of variation in $\boldsymbol{M}$ explained by $\boldsymbol{G}$ (h^2_e); the magnitude of any shared direct effect (τ); and the LD matrix ($\boldsymbol{\Sigma}$, AR(J) structure, $\Sigma_{i,j} = \rho^{|i-j|}$). For all simulations, the expression effects $\boldsymbol{\beta}_E$ and direct effects $\boldsymbol{\theta}$ were sampled independently. More details along with values taken are given in Table 2. For a variable direct effect, τ was set to 0. For a near-constant, shared direct effect, $\tau = 10$. When $\tau = 10$ the direct effects of SNPs on disease are all in the same direction ($\theta_i > 0$ for every SNP i) and the variability in θ in this case is about 1% of the variability when $\tau = 0$. We refer to this situation as "directional pleiotropy.) The process and order for the generation of the simulation data is provided in Table

3. We performed 50K simulation for each combination of parameters (486 different combinations). In each simulation, we generated a new true $\boldsymbol{\beta}_E$ and $\boldsymbol{\theta}$, which are functions of $h_e^2, h_{G\to Y}^2, J$ and τ. Generating this new "true" parameter value better represents the different eQTL and GWAS patterns across different genes, and therefore more resembles a standard TWAS. The procedure detailed in Table 3 is thus repeated 50K times for all combinations of the parameters in Table 2. Type I error (T1E) and power were evaluated at 0.05.

We in addition performed two more sets of simulations to solely assess the T1E. The first scenario, we calculated the empirical T1E under a non-infinitesimal model for both disease and expression (i.e. when only a subset of SNPs were associated with either trait). Under this scenario, SNPs were randomly assigned to be either an eQTL SNP, a disease SNP, or neither. We generated under the scenario where 10% (or 50%) of the SNPs were eQTL SNPs, and 10% (or 50%) of the SNPs were disease SNPs. The LD between these two sets would naturally vary as the SNPs were randomly assigned to their classification. We evaluated the T1E under the same scenarios as above, performing 50K simulations and evaluating at 0.05. We next simulated when the InSIDE condition was violated. We varied the correlation between θ_j and $\beta_{E,j}$ to be 0.125, 0.5, or 0.9. This scenario reflects when there are SNPs nearby not included in the expression model that are in LD with the SNPs in the model.

As $\boldsymbol{\beta}_E$ and $\boldsymbol{\theta}$ were drawn independently, the average bias for the LDA MR-Egger estimates across many independent simulations of $\boldsymbol{\beta}_E$ and $\boldsymbol{\theta}$ should be 0. (This will also be true for the MR-Egger estimates in the absence of LD.) This is analogous to saying the average bias across all the genes in the genome will be zero. However, in practice, we will usually be interested in the test statistic applied to a particular gene. For particular gene with modest *J*, the bias for the LDA MR-Egger need not be 0. To account for this, we next performed a set of 10K simulations with a fixed truth to examine potential bias. We generate one true $\boldsymbol{\beta}_E$ and $\boldsymbol{\theta}$ for each value of *J* that is then held constant for all simulations while we vary the other parameters in Table 2. Therefore, Steps 1 and 5 of Table 3 are only performed once for J=50 or 300. We also compared our estimate of γ to when a regularization factor of 0.1 is added to the diagonal of the LD matrix as mentioned in the beginning of the methods [Pasaniuc and Price 2017].

Application to Breast Cancer GWAS Summary Data

We next applied TWAS, LDA MR, and LDA-MR Egger analyses applied to a breast cancer GWAS. We previously conducted a separate breast cancer TWAS using a different approach to build expression weights and including both validation of predicted expression and functional follow up of significant genes [Wu, et al. 2017]. Here our focus is to compare different analysis approaches where each analysis uses the same set of simple expression weights and the same GWAS summary statistics. These expression weights are different from that of the Wu et al paper [Wu, et al. 2017]. The marginal GWAS summary statistics were from a recent GWAS on breast cancer within women of European descent [Michailidou, et al. 2017]. SNP data was meta-analyzed from 13 platforms: OncoArray [Amos, et al. 2017], iCOGS, and eleven other GWAS. More details on the QC procedure can be found in Michailidou et al [Michailidou, et al. 2017]. After QC, the study consisted of 11.8 million SNPs, with 105,974 controls and 122,977 cases. The GWAS estimates were calculated on the non-standardized minor allele counts, and therefore were transformed using the minor allele frequency.

Expression weights were calculated from GTEx along with LD information in breast tissue in an overall sample of 183 individuals [Consortium 2013]. We restricted our analysis to the set of transcripts that were deemed heritable using GCTA [Yang, et al. 2011] and examined SNPs within 500kb of the gene boundary. The expression weights were calculated on standardize minor allele counts of SNPs (mean zero, variance one) and were conditionally estimated using the BSLMM approach [Zhou, et al. 2013]. A gene was deemed heritable if the GCTA p-value for each tissue from GTEx was less than the Bonferroni threshold of 0.05 (after adjusting for 27,945 tests). We were left with 683 transcripts to analyze.

We analyze the Breast Cancer data for these genes using the TWAS, LDA MR, and LDA MR-Egger. We did not examine the MR and MR-Egger as we were testing for cis-signals as opposed to genome wide and therefore the SNPs were in LD. We took the overlap of the $\widehat{\boldsymbol{\beta}}_G$ SNPs from the GWAS with the available SNP correlations from GTEx. If the effects were estimated with respect to a different reference allele, we reversed the sign for that eQTL effect estimate. In total, the 683 genes corresponded to 191,583 unique SNPs.

Results

Simulation Study

First, we examined the type I error rate in simulations and observed that when there is little LD and no direct effect of the SNPs, all of the approaches have the correct type I error (Figure 1, Sup Figure 1). When there is little LD and the direct effect is variable, the MR and MR-Egger approaches have modestly inflated Type I Error rates. When there is low LD and there is directional pleiotropy, only the LDA MR-Egger has correct type I error. If the SNPs are in high LD (bottom row of Figure 1) and there is no direct effect, the MR and the MR-Egger have inflated type I error due to misspecification of the variance. When there is a variable direct effect, all four approaches have inflated type I error. Finally, when there is strong linkage and directional pleiotropy, only the LDA MR-Egger has the correct type I error.

We next examined the power when there is little to no LD between the SNPs and no direct effect (Figure 2). Under this situation, all four approaches had correct type I error (Figure 1 top left plot, Sup Figure 1 top left plot) and are valid tests. Regardless of the magnitude of the effect of **M** on **Y**, we see similar power for all four methods. There is slightly smaller power for the LDA MR-Egger compared to the LDA MR or MR when the SNPs only explain 20% of the variation in the gene expression, but once the SNPs explain 50% of the variation in $\boldsymbol{M}$, all approaches have approximately equal power. In each individual plot, we see that as the SNPs explain more of the variation (and are thus better instruments), we have an increase in power. It has been previously reported that the MR Egger regression can have notably lower power than MR [Bowden, et al. 2015]. We did not observe large differences in power, but we consider situations with more SNPs in the genetic instrument for the intermediate factor and largr GWAS sample sizes. (For example, simulations in Bowden et al. (2015) consider 25 SNPs up to 1,000 subjects in the GWAS.) We performed additional simulations assuming only 5 SNPs were included in the genetic instrument; in this case, we also observed lower power for the LDA MR Egger method relative to the LDA MR method.

The decrease in power from J=50 to J=300 is due to an increase in the noise to signal ratio when predicting expression levels using SNPs. For J=50, a larger proportion of the variation explained by SNPs is shared by each individual SNP leading to more precise estimates. When J=300, a smaller proportion of that same amount of variation is explained by each SNP in a larger set of SNPs. Assuming the sample size in the reference panel used to estimate $\widehat{\boldsymbol{\beta}}_E$ is the

same, the sampling error in the SNP-specific estimates $\hat{\beta}_{E,j}$ is the same for J=50 and J=300. We have held the proportion of variance explained by the SNPs constant, while increasing the noise due to sampling error (as we are using estimated expression effects from 300 SNPs rather than 50).

When there is low LD and directional pleiotropy (Table 4), then the MR-Egger approach has less power than the LDA MR-Egger test to detect an association when J=50. This may be due to LDA MR-Egger regression gaining some information by accounting for the weak LD structure. If J=300, there is a decrease in power compared to when J=50 for both of these methods regardless of the presence of a direct. At J=300, the LDA MR-Egger has slightly higher power than the MR-Egger. When the SNPs explain 50% of the variation in $\boldsymbol{M}$, both methods have power greater than 80% regardless of the effect of $\boldsymbol{M}$ on $\boldsymbol{Y}$ (Table 4). We did not report the power of the MR or LDA MR, as they did not have proper type I error when there is directional pleiotropy (Figure 1 and Supplement Figure 1).

Finally, we examine the power when there is strong LD amongst the SNPs. We do not assess the MR or MR-Egger for this scenario as they do not have correct type I error for correlated SNPs. When J=50, and there is no direct effect, the LDA MR has more power than the LDA MR-Egger, though the difference in power is less pronounced when h_E^2=0.5 and $\boldsymbol{M}$ has a strong effect on $\boldsymbol{Y}$. When J=300, the two methods have comparable power (Table 4), with the LDA MR-Egger having slightly less power than the LDA MR. When there is a direct effect, the LDA MR-Egger has approximately equal power as when there is no direct effect.

When J=50, $h_E^2 = 0.5$, and there is a small mediated effect, then LDA MR-Egger regression has more power when there is strong LD versus low LD (power of 0.914 versus 0.784). When J=300, there is a smaller drop in power for small LD to strong LD, with the LDA MR-Egger power going from 0.869 to 0.840. The LDA MR approach has equal power regardless of whether the SNPs are strongly correlated vs weakly correlated.

The results for the non-infinitesimal simulations are provided in Supplementary Figures 2-5. For the most part, these results mirror what was observed in the standard T1E analysis. The only difference was when J=50 and there was little LD (Supplement Figure 2). When there is a variable direct effect and the SNPs explain an adequate amount of the variation in gene

expression, we see all four approaches are conservative. In this case, where a small number of SNPs have relatively large effects on expression, the residual errors in the regression of GWAS effects on eQTL weights may be increased, which in turn will deflate the test statistics.

When the InSIDE conditions are violated, we see that regardless of if the direct effect are near-constant or variable around 0 we have massively inflated T1E (Figure 3 and variable direct effect Supplement Figure 6). The results for when J=300 are given in the supplement (Supplement Figures 7 and 8).

We next examine the bias of our estimates (Figure 4). We here show the results when there is strong LD (Figure 4). The results for low are in LD Supplement Figure 9. In Figure 4, the first two rows show the results when J=50, and the next two rows when J=300. The first column shows when there is no direct effect, the second column when there is a variable direct effect, and the third column when there is directional pleiotropy. The first and third row are when there is no effect of the mediator on the outcome ($\gamma = 0$) and the second and fourth row show when there is a large effect ($\gamma \neq 0$). Here we have fixed $\boldsymbol{\beta}_E$ and $\boldsymbol{\theta}$ effects for all simulations.

As expected based on the section on bias of estimates sections, when $\gamma = 0$ and $\boldsymbol{\theta} = \mathbf{0}$**,** all of the estimates are unbiased. When there is an effect on the outcome and no direct effect, we see that the non LDA aware approaches converge faster to the truth than the LD aware methods. While the non LDA converge faster, recall that they misspecify the variance and lead to improper inference. We also see attenuation bias when $\gamma \neq 0$, with estimates improving as the SNPs become better instruments. The attenuation bias is larger for J=300 relative to J=50, for the same reason that we saw a decrease in power from J=50 to J=300: an increase in the noise to signal ratio. When there is a variable direct effect, all of the approaches are biased (second column). Our empirical bias result depends on the particular values of $\boldsymbol{\beta}_E$ and $\boldsymbol{\theta}$. As described in the methods ("Bias"), all of the estimates for the mediated effect of gene expression include a weighted covariance between $\boldsymbol{\beta}_E$ and $\boldsymbol{\theta}$. Even if the average across all genes for this covariance is zero, for any particular gene it is non-zero, and can be large. If the $\boldsymbol{\beta}_E$ and $\boldsymbol{\theta}$ for each gene can be thought of as independent draws from their respective distributions, then the average absolute magnitude of the bias term is a decreasing function of *J*. This is why we see smaller bias when *J*=300. Finally, when there is directional pleiotropy, only the LDA MR-Egger is unbiased when $\gamma = 0$ or J=50. When J=300 and $\gamma \neq 0$, we see the attenuation bias in the LDA-MR Egger, with

it biased downward from the null. We saw little difference between the LDA methods based on adding a regularization constant of 0.1 to the diagonal or not of the LD matrix (Supplement Figure 10-11).

Application to Breast Cancer GWAS Summary Data

Of the 683 genes tested, 79 were called significant ($p< 7.32*10^{-5}$) by at least one approach (TWAS, LDA MR or LDA MR-Egger, Supplementary Tables). Comparing the TWAS vs the LDA MR-Egger (Figure 5A), there were 12 genes that were significant by the TWAS but not by LDA MR-Egger, 26 genes that were by LDA MR-Egger and not TWAS, and 20 that were called by both. With the LDA MR and the LDA MR-Egger, there was much more agreement due to the same weight matrix being used, but still the LDA MR called 8 genes as significant that the LDA-MR Egger did not (Figure 5B). Thirty-eight gene transcripts were called significant by both the LDA-MR and the LDA-MR Egger methods (Supplement Tables 1 and 3). There were 19 genes called significant by all three approaches. A detailed list of which gene transcripts were found significant by which method is provided in Supplementary Tables 2 through 8. Examining the spearman correlation between the p-values, LDA MR and LDA MR–Egger had an r^2 of 0.45, LDA MR and TWAS of .51, and LDA MR-Egger and TWAS of 0.31. The kappa statistic for calling a gene transcript significant between LDA MR and LDA MR-Egger was 0.63, between LDA MR and TWAS was 0.55, and between LDA MR-Egger and TWAS was 0.48. We note that although the TWAS and LDA MR statistics are equivalent when the GWAS effect estimates have constant variance (see methods), in practice the GWAS effect estimates will differ across SNPs (e.g. due to sample size differences), leading to the differences in test statistics we see here. The Pearson correlation squared between TWAS and LDA MR test-statistics was 0.72 (Supplement Figure 12).

A gene that we will highlight that was called significant by the TWAS was SET Domain Containing 9 (*SETD9*), with a p-value of 2.47e-23 at cytoband 5q11.2. The p-value for this gene transcript for the LDA MR and the LDA MR-Egger was 2.97e-09 and 0.41 respectively. This is potentially a TWAS and LDA-MR false-positive as the LDA MR-Egger was not associated with gene. A fine mapping analysis of this locus found four functional candidate SNPs in a sample of approximately 100K women of European descent [Glubb, et al. 2015]. These four candidate

functional SNPs were associated with an increase in activity of *MAP3K1* (Mitogen-Activated Protein Kinase 1), another gene at this locus. *SETD9* was ruled as not the gene of interest as it had no association with these four candidate SNPs. *MAP3K1* is located 94 kbp from *SETD9*. (*MAP3K1* did not pass our cis-heritable tissue threshold and was not included in the analysis.)

Discussion

In this work we connect TWAS with Mendelian Randomization (MR) and examine the performance of the LD aware MR-Egger regression in the context of causal gene expression identification. In contrast to TWAS that mainly focused on novel risk region identification, here we investigate causal gene identification at a given risk region, a more difficult problem; for example, traditional TWAS does not aim to disentangle between two genes that show significant associations at the same risk region (e.g., due to LD-tagging between SNPs in the expression prediction models for the two genes). We show that the standard TWAS method is a particular case of MR. We investigate a variant of MR that allows for LD (LD aware MR-Egger) that can account for LD and direct effects under the scenario when the variability in direct effects across SNPs is small and the InSIDE condition holds. The LDA MR, the summary MR and the MR-Egger are not proper estimates under these scenarios as they either do not account for: the direct effect (LDA MR, MR) or the correlation between SNPs (MR and MR-Egger). For scenarios where the LDA MR is a valid test, the LDA MR-Egger has slightly lower power in our simulations. We also provide a one-to-one relationship between TWAS and the traditional LDA-MR when the standard errors of the marginal GWAS effects are all equal. In our real data application, the three approaches (TWAS, LDA MR and LDA MR-Egger) called different genes significant, with the LDA MR-Egger potentially correctly calling some results as null that the TWAS did not. While we focused on the case of gene expression data, the LDA MR-Egger can be extended to an arbitrary mediator when the instruments are correlated.

The majority of our simulations assumed an "infinitesimal" genetic architecture for both the gene expression and outcome phenotypes. Departures from this model—for example, if gene expression is causally influenced by only one or a small proportion of local SNPs not included in the model—can affect the performance of the tests. We saw this in our simulations where the InSIDE condition was violated. Future work could consider the impact of local genetic architecture on these tests. We mainly focused on the case where there may be pleiotropic direct

effects on the outcome, but these are not systematically related to the SNPs' effects on the mediator. This is reflected in our simulations when we draw $\boldsymbol{\beta}_E$ and $\boldsymbol{\theta}$ independently. If there is a systematic relationship between $\boldsymbol{\beta}_E$ and $\boldsymbol{\theta}$, then none of the methods we have discussed here will provide unbiased estimates of γ or valid tests of $\gamma = 0$ (Figure 3). There could be a systematic relationship between $\boldsymbol{\beta}_E$ and $\boldsymbol{\theta}$ even when $\gamma = 0$ if, for example, the SNPs influence a third (unobserved) trait, which in turn influences both the mediator and outcome. This is arguably unlikely to be true when the mediator is expression of a gene and the genetic instrument consists of cis SNPs, except in the case of co-localization. Evaluating the causal effect of the mediator on outcome when the instrument and direct effects are systematically related is not possible without additional information on the mediator-outcome relationship: large samples with data on genetic factors, the mediator, outcome and possible confounders will likely be needed.

In practice, we note that meaningful biological interpretation of the magnitude of γ may be difficult, due to QC and pre-processing (scaling) of the data. (The sign of γ, however, may contain useful information regarding the directionality of expression effects on disease.) We are also faced with the issue of attenuation bias, as we have estimates of the true eQTL parameters. Despite this, the LDA MR-Egger is still a valid test for the effect of the gene on outcome if the required assumptions are met. To reiterate these assumptions: they require that the variability in the direct effect of the SNPs on disease is relatively small and that the mechanism that the direct effect acts through is independent of the eQTL effect (InSIDE condition). The LDA MR-Egger can still however help to highlight a region of interest. Practitioners should follow up any analyses with a deep literature search paired with examining functional information such as pathways, enhancers, and other tissue expression. Whereas traditional epidemiological MR studies need to be certain of the causal pathways, the goal of these methods is to identify target genes whose expression levels are likely to influence disease risk. These genes represent candidates for functional experiments in model systems to investigate the effects of perturbing gene expression.

In summary, we have examined the performance of various summary statistics approaches and how they compare to the LDA MR-Egger. The LDA MR-Egger approach can be utilized only under the following situations: 1) No Direct Effect, 2) There is limited variability in

direct effects of the SNPs on the outcome and this effect is independent of the eQTL effect. If we are not in one of those scenarios, the LDA MR-Egger test for the gene expression's effect on outcome will not be valid—although neither will any of the other tests we considered This work provides guidance on the interpretation of TWAS tests and suggests that LDA MR-Egger regression may be a useful sensitivity analysis in situations where false positives due to colocalization are a concern.

Acknowledgments

PK and BP were supported by U.S. National Institute of Health grants R01 HG009120. PK was in addition funded by P01 CA134294. RB was supported by the U.S. National Cancer Institute grant R35 CA197449. AG was supported by U.S. National Institute of Health grants R01 GM105857. The breast cancer genome-wide association analyses were supported by the Government of Canada through Genome Canada and the Canadian Institutes of Health Research, the 'Ministère de l'Économie, de la Science et de l'Innovation du Québec' through Genome Québec and grant PSR-SIIRI-701, The National Institutes of Health (U19 CA148065, X01HG007492), Cancer Research UK (C1287/A10118, C1287/A16563, C1287/A10710) and The European Union (HEALTH-F2-2009-223175 and H2020 633784 and 634935). All studies and funders are listed in Michailidou et al (2017). The computations in this paper were run on the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University.

References

Amos CI, Dennis J, Wang Z, Byun J, Schumacher FR, Gayther SA, Casey G, Hunter DJ, Sellers TA, Gruber SB and others. 2017. The OncoArray Consortium: A Network for Understanding the Genetic Architecture of Common Cancers. Cancer Epidemiol Biomarkers Prev 26(1):126-135.

Barbeira A, Shah KP, Torres JM, Wheeler HE, Torstenson ES, Edwards T, Garcia T, Bell GI, Nicolae D, Cox NJ and others. 2016. MetaXcan: Summary Statistics Based Gene-Level Association Method Infers Accurate PrediXcan Results. bioRxiv.

Bowden J, Davey Smith G, Burgess S. 2015. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 44(2):512-25.

Burgess S, Dudbridge F, Thompson SG. 2016. Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. Stat Med 35(11):1880-906.

Burgess S, Thompson SG. 2017. Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol 32(5):377-389.

Consortium GT. 2013. The Genotype-Tissue Expression (GTEx) project. Nat Genet 45(6):580-5.

Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, Eyler AE, Denny JC, Consortium GT, Nicolae DL and others. 2015. A gene-based association method for mapping traits using reference transcriptome data. Nat Genet 47(9):1091-8.

Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V. 2014. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet 10(5):e1004383.

Glubb DM, Maranian MJ, Michailidou K, Pooley KA, Meyer KB, Kar S, Carlebur S, O'Reilly M, Betts JA, Hillman KM and others. 2015. Fine-scale mapping of the 5q11.2 breast cancer locus reveals at least three independent risk variants regulating MAP3K1. Am J Hum Genet 96(1):5-20.

Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, Jansen R, de Geus EJ, Boomsma DI, Wright FA and others. 2016. Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet 48(3):245-52.

Han B, Kang HM, Eskin E. 2009. Rapid and accurate multiple testing correction and power estimation for millions of correlated markers. PLoS Genet 5(4):e1000456.

Hormozdiari F, van de Bunt M, Segre AV, Li X, Joo JWJ, Bilow M, Sul JH, Sankararaman S, Pasaniuc B, Eskin E. 2016. Colocalization of GWAS and eQTL Signals Detects Target Genes. Am J Hum Genet 99(6):1245-1260.

Johnson T. 2011. Efficient calculation for multi-SNP genetic risk scores. Technical Report, Queen Mary University of London.

Mancuso N, Kichaev G, Shi H, Freund M, Giambartolomei C, Gusev A, Pasaniuc B. 2017a. Probabilistic fine-mapping of transcriptome-wide association studies. bioRxiv.

Mancuso N, Shi H, Goddard P, Kichaev G, Gusev A, Pasaniuc B. 2017b. Integrating Gene Expression with Summary Association Statistics to Identify Genes Associated with 30 Complex Traits. Am J Hum Genet 100(3):473-487.

Michailidou K, Lindstrom S, Dennis J, Beesley J, Hui S, Kar S, Lemacon A, Soucy P, Glubb D, Rostamianfar A and others. 2017. Association analysis identifies 65 new breast cancer risk loci. Nature.

Pasaniuc B, Price AL. 2017. Dissecting the genetics of complex traits using summary association statistics. Nat Rev Genet 18(2):117-127.

Shi H, Kichaev G, Pasaniuc B. 2016. Contrasting the Genetic Architecture of 30 Complex Traits from Summary Association Data. Am J Hum Genet 99(1):139-53.

Valeri L, VanderWeele TJ. 2013. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods 18(2):137-50.

Wainberg M, Sinnott-Armstrong N, Knowles D, Golan D, Ermel R, Ruusalepp A, Quertermous T, Hao K, Bjorkegren JLM, Rivas MA and others. 2017. Vulnerabilities of transcriptome-wide association studies. bioRxiv.

Wallace C. 2013. Statistical testing of shared genetic control for potentially related traits. Genet Epidemiol 37(8):802-13.

Wu L, Shi W, Long J, others a. 2017. Identification of novel susceptibility loci and genes for breast cancer risk: A large transcriptome-wide association study in nearly 230,000 women of European descent. Submitted.

Yang J, Lee SH, Goddard ME, Visscher PM. 2011. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 88(1):76-82.

Zhou X, Carbonetto P, Stephens M. 2013. Polygenic modeling with bayesian sparse linear mixed models. PLoS Genet 9(2):e1003264.

Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, Montgomery GW, Goddard ME, Wray NR, Visscher PM and others. 2016. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 48(5):481-7.

Figure Legends

Figure 1: Type I error when J=50. Each bar represents results over $5x10^4$ simulations. Evaluated at $\alpha = 0.05$. First panel represent when low LD (plots with A). Second panel represents when strong LD (plots with B). From left to right correspond to: no direct effect, variable direct effects with mean 0 across SNPs, and direct effects with mean >0 across SNPs and small variability (directional pleiotropy). When there is a direct effect, the SNPs explain 1% of the variation in the outcome ($h^2_{G\rightarrow Y}$=0.01).

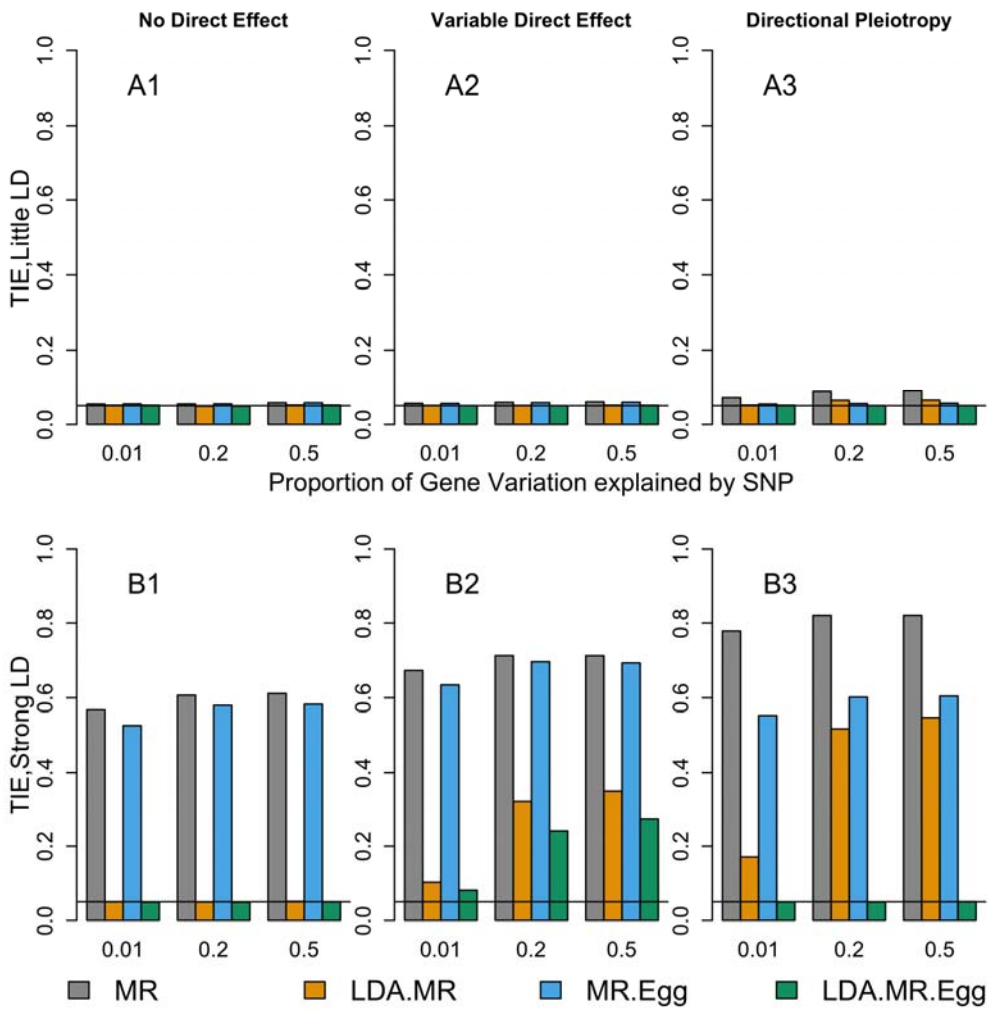


Figure 2: Power when little to no LD and J=50, 300. Power results when there is little to no LD and no direct effect. Each bar represents results over $5x10^4$ simulations. Evaluated at $\alpha = 0.05$. First row represents J=50 and second row when J=300. From left to right: when $\gamma^2 = 0.005$ and $\gamma^2 = 0.01$.

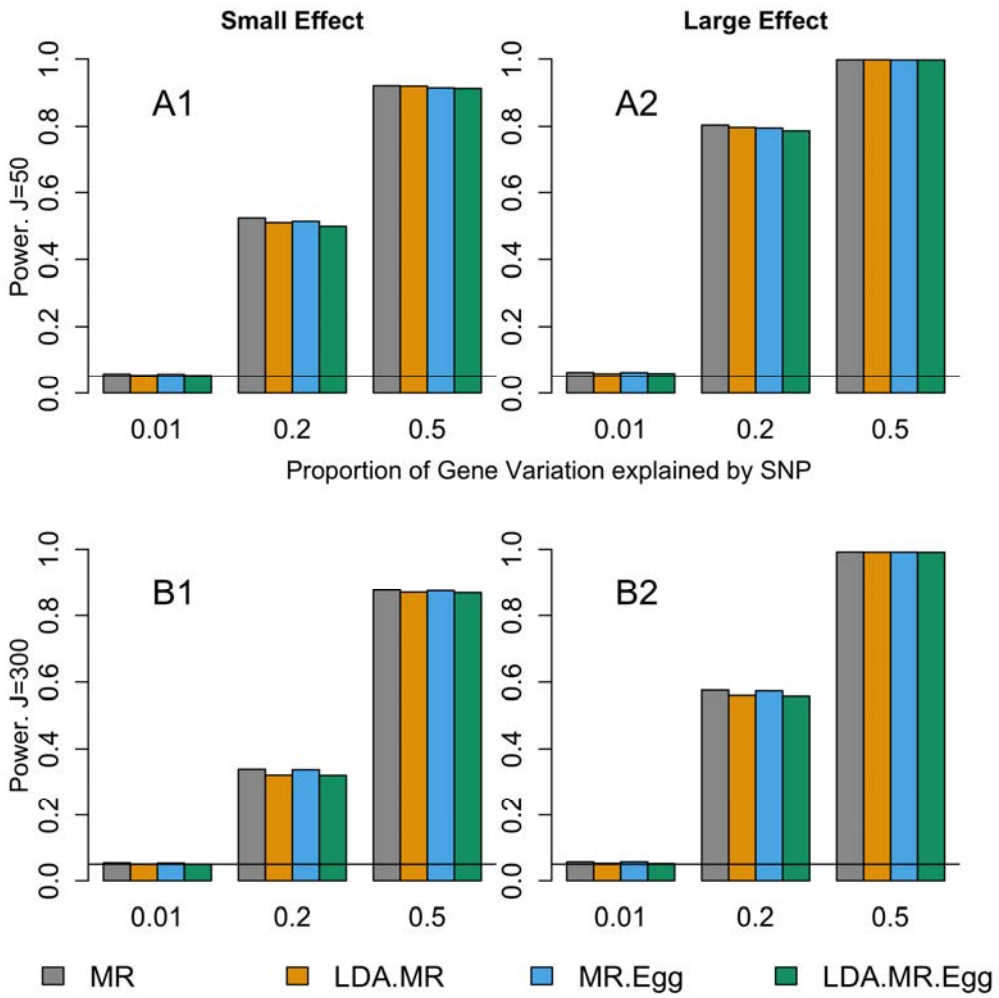


Figure 3: Type I error when J=50, InSIDE condition violated, and there is directional pleiotropy. Each bar represents results over $5x10^4$ simulations. Evaluated at $\alpha = 0.05$. First panel represent when low LD (plots with A). Second panel represents when strong LD (plots with B). From left to right correspond to: correlation between θ_j and β_{EJ} is 0.125, 0.5, or 0.9.

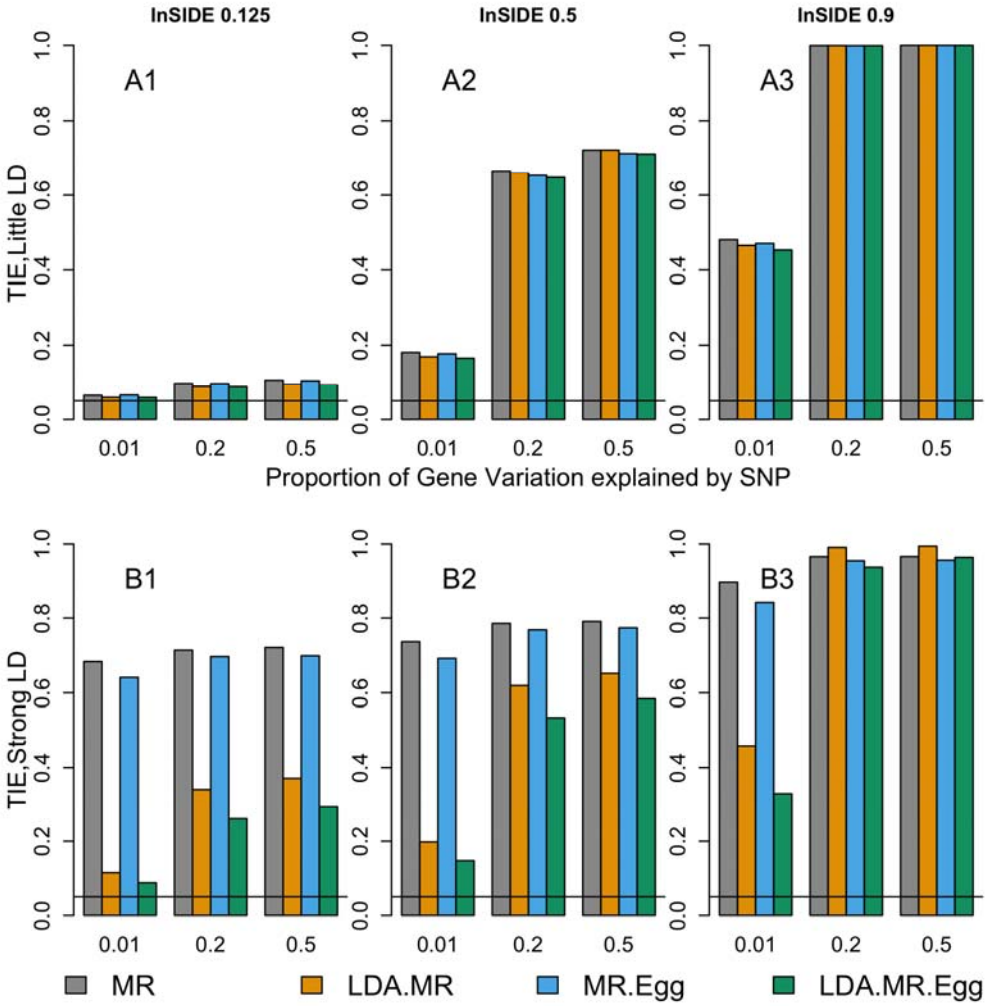


Figure 4: Bias when strong LD for J=50, 300. Bias plots for when there is strong LD in the SNP set. First row corresponds to J=50, $\gamma = 0$ (plots with A). Second panel (plots with B) when J = 50 and $\gamma^2 = 0.01$. Third panel (plots with C) when J = 300 and $\gamma = 0$. Final panel (plots with D) J = 300 and $\gamma^2 = 0.01$. From left to right: no direct effect, variable direct effects with mean 0 across SNPs, and directional pleiotropy. When there is a direct effect, the SNPs explain 1% of the variation in the outcome ($h^2_{G\rightarrow Y}$=0.01).

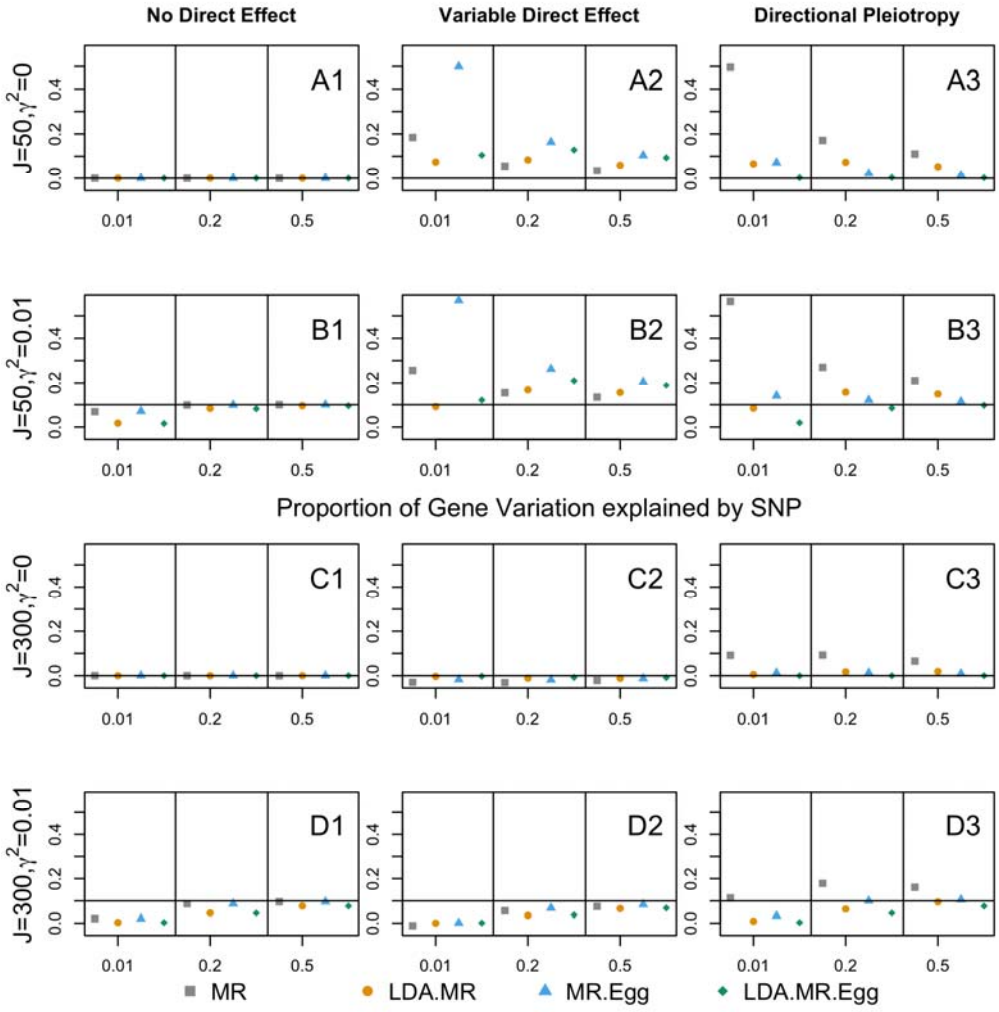
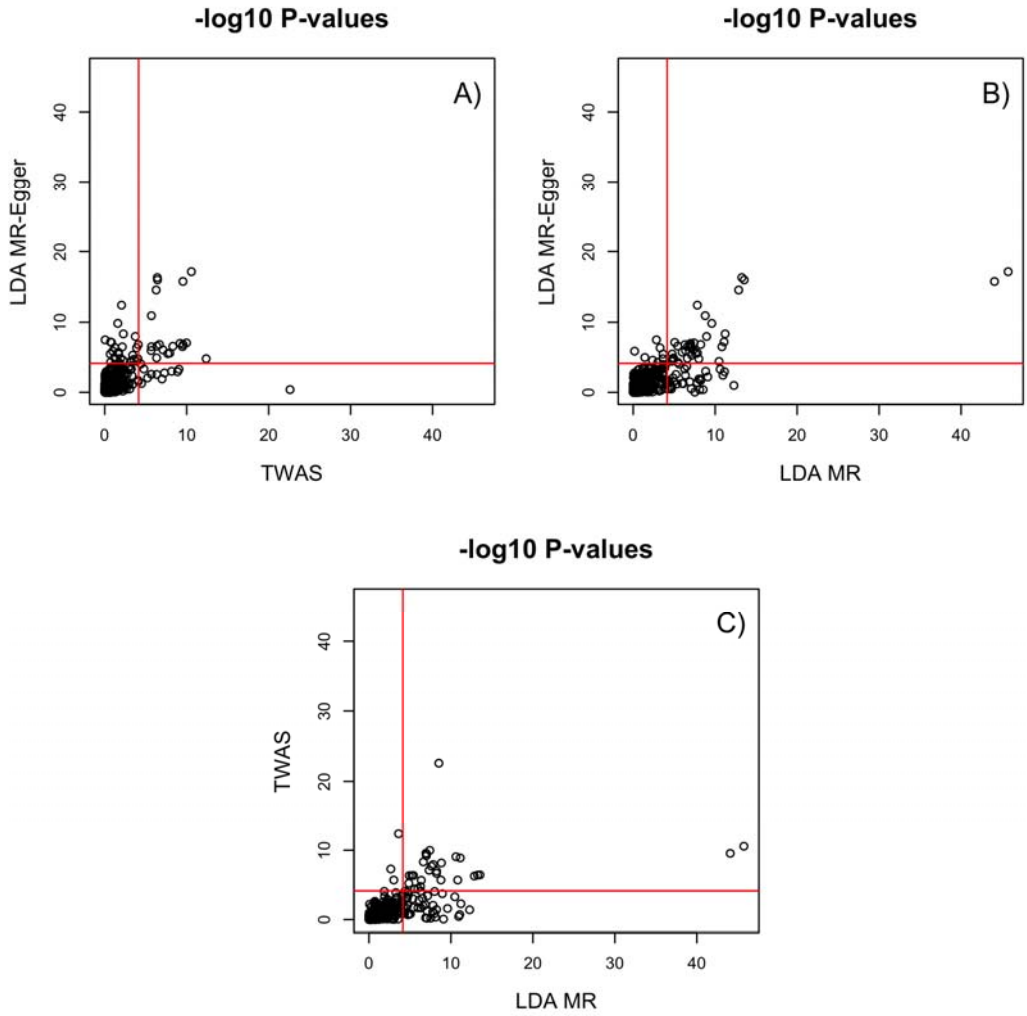


Figure 5: Comparing –log 10 p-values. Shows for 683 genes between LDA MR Egger and TWAS (A), LDA MR Egger and LDA MR (B), and LDA MR and TWAS (C). Red line is the Bonferroni cutoff of $-\log_{10}(.05/683)$.



Table

Table 1: Commonly used approaches for testing using summary statistics. Details four common Summary Mendelian Randomization approaches and the TWAS for testing for an association between gene expression and outcome through GWAS (*J* is the number of SNPs in the loci).

Method	Adjusts for Direct Effect	Accounts for LD	Distribution
TWAS	No	Yes	

MR	No	No	$Z_{MR} \sim t(df = J - 1)$
MR-Egger	Yes	No	$Z_{MRE} \sim t(df = J - 2)$
LDA MR	No	Yes	$Z_{LDMR} \sim t(df = J - 1)$
LDA MR-Egger	Yes	Yes	$Z_{LDMRE} \sim t(df = J - 2)$

Table 2: Simulation parameters that were modified. Parameters that were varied for all simulations performed. A description of the parameter and potential values could be taken are given. All combinations of parameters were examined.

Parameter	Definition	Values taken
h_e^2	Proportion of variability in E explained by G	0.01, 0.20, 0.50
$h_{E\to Y}^2$	Proportion of Variability in Y explained by E	0, 0.005, 0.01
$h_{G\to Y}^2$	Proportion of Variability in Y explained by G	0, 0.005, 0.01
N	Sample Size of GWAS	5000
N_E	Sample Size of EQTL	1000
ρ	Correlation between SNPs. AR structure	0.125, 0.9
J	Number of SNPs in the Loci	50,300
τ	Strength of Pleiotropic Effects	0,10
$\mathbf{L}$	Cholesky Decomposition of $\boldsymbol{\Sigma}$	Function of ρ

Table 3: Procedure for generating simulation study. Gives the steps in order for generating the type I error and power simulations detailed in the paper and how the parameters were generated.

Step	Procedure	Mathematically
1	Generate true eQTL	$\boldsymbol{\beta}_E \sim N_J(0, \boldsymbol{I}_J)$
2	Set the proportion of variability in expression due to SNPs	$\sigma_E = \sqrt{\frac{h_e^2}{\boldsymbol{\beta}_E^T \boldsymbol{\Sigma} \boldsymbol{\beta}_E}}$
3	Rescale the eQTL effects	$\boldsymbol{\beta}_E = \sigma_E \boldsymbol{\beta}_E$
4	Set Expression to outcome Effect	$\gamma = \sqrt{h_{E\to Y}^2}$
5	Generate potential Direct Effect	$\boldsymbol{\theta} = \boldsymbol{e} + \tau;\ \boldsymbol{e} \sim N_J(0, \boldsymbol{I}_J)$
6	Set the proportion of variability in outcome due directly to SNPs	$\sigma = \sqrt{\frac{h_{G\to Y}^2}{\boldsymbol{\theta}^T \boldsymbol{\Sigma} \boldsymbol{\theta}}}$
7	Rescale the Direct Effects	$\boldsymbol{\theta} = \sigma * \boldsymbol{\theta}$
8	Generate the GWAS effects	$\boldsymbol{\beta}_G = \boldsymbol{\theta} + \gamma \boldsymbol{\beta}_E$
9	Generate Observed eQTL Values	$\widehat{\boldsymbol{\beta}}_E^* \sim \boldsymbol{\Sigma} \boldsymbol{\beta}_E + \boldsymbol{L}^T \boldsymbol{\epsilon}_E;\ \boldsymbol{\epsilon}_E \sim N_J\left(0,\ \frac{1 - h_E^2}{N_E} \boldsymbol{I}_J\right)$
10	Generate Observed GWAS values	$\widehat{\boldsymbol{\beta}}_G^* \sim \boldsymbol{\Sigma} \boldsymbol{\beta}_G + \boldsymbol{L}^T \boldsymbol{\epsilon}_G;\ \boldsymbol{\epsilon}_G \sim N_J\left(0, \boldsymbol{I}_J\ \frac{1 - h_E^2 * h_{E\to Y}^2 - h_{G\to Y}^2}{N}\right)$

Table 4: Power results from simulation study. Details the power results at an α level of 0.05 under varying levels of LD, number of SNPs in the loci, and presence of a direct effect. TWAS not shown as is equivalent to the LDA MR under are simulation procedure.

LD	Number of SNPS	% Variation E explained by SNPs	Method	No Direct effect		Directional Pleiotropy	
		Strength of Mediated Effect		Small Effect	Large Effect	Small Effect	Large Effect
			MR	0.058	0.063	----	----
			LDA MR	0.054	0.058	----	----
		0.01	MR-Egger	0.057	0.062	0.055	0.057
			LDA MR - Egger	0.054	0.059	0.055	0.059
			MR	0.524	0.804	----	----
			LDA MR	0.510	0.797	----	----
	J=50	0.2	MR-Egger	0.514	0.795	0.302	0.508
			LDA MR - Egger	0.499	0.787	0.498	0.788
			MR	0.922	0.997	----	----
			LDA MR	0.921	0.997	----	----
		0.5	MR-Egger	0.915	0.997	0.658	0.911
Small LD			LDA MR -	0.914	0.997	0.915	0.997

		Egger				
		MR	0.054	0.057	----	----
		LDA MR	0.051	0.053	----	----
	0.01	MR-Egger	0.054	0.057	0.052	0.051
		LDA MR - Egger	0.050	0.052	0.052	0.051
		MR	0.339	0.576	----	----
		LDA MR	0.321	0.560	----	----
J=300	0.2	MR-Egger	0.338	0.573	0.288	0.501
		LDA MR - Egger	0.321	0.557	0.325	0.562
		MR	0.877	0.992	----	----
		LDA MR	0.870	0.992	----	----
	0.5	MR-Egger	0.875	0.992	0.812	0.980
		LDA MR - Egger	0.869	0.991	0.869	0.992
	0.01	LDA MR	0.054	0.059	----	----
		LDA MR - Egger	0.052	0.055	0.053	0.055

	J=50	0.2	LDA MR	0.513	0.798	----	----
			LDA MR - Egger	0.389	0.629	0.389	0.637
		0.5	LDA MR	0.919	0.998	----	----
Large LD			LDA MR - Egger	0.784	0.943	0.788	0.941
		0.01	LDA MR	0.049	0.049	----	----
			LDA MR - Egger	0.050	0.049	0.051	0.051
	J=300	0.2	LDA MR	0.318	0.560	----	----
			LDA MR - Egger	0.294	0.523	0.300	0.524
		0.5	LDA MR	0.869	0.992	----	----
			LDA MR - Egger	0.840	0.985	0.842	0.986