

Testing Genetic Pleiotropy with GWAS Summary Statistics for Marginal and Conditional Analyses

Yangqing Deng and Wei Pan¹

Division of Biostatistics, University of Minnesota, Minneapolis, Minnesota 55455

ABSTRACT There is growing interest in testing genetic pleiotropy, which is when a single genetic variant influences multiple traits. Several methods have been proposed; however, these methods have some limitations. First, all the proposed methods are based on the use of individual-level genotype and phenotype data; in contrast, for logistical, and other, reasons, summary statistics of univariate SNP-trait associations are typically only available based on meta- or mega-analyzed large genome-wide association study (GWAS) data. Second, existing tests are based on marginal pleiotropy, which cannot distinguish between direct and indirect associations of a single genetic variant with multiple traits due to correlations among the traits. Hence, it is useful to consider conditional analysis, in which a subset of traits is adjusted for another subset of traits. For example, in spite of substantial lowering of low-density lipoprotein cholesterol (LDL) with statin therapy, some patients still maintain high residual cardiovascular risk, and, for these patients, it might be helpful to reduce their triglyceride (TG) level. For this purpose, in order to identify new therapeutic targets, it would be useful to identify genetic variants with pleiotropic effects on LDL and TG after adjusting the latter for LDL; otherwise, a pleiotropic effect of a genetic variant detected by a marginal model could simply be due to its association with LDL only, given the well-known correlation between the two types of lipids. Here, we develop a new pleiotropy testing procedure based only on GWAS summary statistics that can be applied for both marginal analysis and conditional analysis. Although the main technical development is based on published union-intersection testing methods, care is needed in specifying conditional models to avoid invalid statistical estimation and inference. In addition to the previously used likelihood ratio test, we also propose using generalized estimating equations under the working independence model for robust inference. We provide numerical examples based on both simulated and real data, including two large lipid GWAS summary association datasets based on ~100,000 and ~189,000 samples, respectively, to demonstrate the difference between marginal and conditional analyses, as well as the effectiveness of our new approach.

KEYWORDS GEE; likelihood ratio test; multiple-trait association testing; structural equation models; union-intersection test; Wald test

THERE has been a growing interest in testing genetic pleiotropy, which is when a single genetic variant (or gene) influences multiple traits (Solovieff *et al.* 2013). As pointed out by Schaid *et al.* (2016), detecting pleiotropy may shed light on the underlying biological mechanism of a genetic association, with important implications for therapeutic development, such as drug repurposing. Testing pleiotropy can be also useful in checking an assumption (*i.e.*, no pleiotropic effects of a genetic instrument variable on both a risk factor/exposure and an outcome) imposed by Mendelian randomization for causal inference. There is accumulating empirical evidence to support pervasive pleiotropy for com-

plex diseases and traits (Cotsapas *et al.* 2011; Q. Wang *et al.* 2015). Leveraging pleiotropy in GWAS with multiple traits may boost statistical power in detecting genetic associations (Chung *et al.* 2014) as well as improving genetic risk prediction (Li *et al.* 2014). Most existing methods for analysis of multiple traits test on a global null hypothesis that none of the traits is associated with a genetic variant, which, if rejected, cannot tell whether the genetic variant is associated with only one or with two or more traits (*e.g.*, Y. Wang *et al.* 2015; Kim *et al.* 2016; Z. Wang *et al.* 2016 and references therein). In pleiotropy testing, the target of interest is not the global null hypothesis, but rather that the genetic variant is associated with no more than one trait. A simple approach would first conduct a univariate association test on each trait separately, then check whether two or more *P*-values are significant after a proper multiple testing adjustment. As shown in a later example, this naïve approach is often too

Copyright © 2017 by the Genetics Society of America

doi: <https://doi.org/10.1534/genetics.117.300347>

Manuscript received July 4, 2017; accepted for publication September 29, 2017; published Early Online October 2, 2017.

¹Corresponding author: Division of Biostatistics, University of Minnesota, 420 Delaware St SE, Minneapolis, MN 55455. E-mail: weip@biostat.umn.edu

low powered due to the univariate nature of this test. Accordingly, some methods have been developed to test pleiotropy. Cotsapas *et al.* (2011) developed a cross-phenotype meta-analysis (CPMA) method to test pleiotropy using summary statistics, which seems straightforward, but requires that we already know of the existence of one significant association, which may not be practical. In addition, it imposes a strong parametric assumption on the distribution of the marginal association P -values under the alternative hypothesis, which may lead to loss of power when the assumption does not hold. Alternatively, other methods explore which of the multiple traits are associated (Stephens 2013; Majumdar *et al.* 2016). Schaid *et al.* (2016) not only nicely surveyed existing methods, but also proposed a new and rigorous pleiotropy test based on the intersection-union testing approach (Berger 1997) with individual-level genotype and phenotype data. However, Schaid *et al.* (2016) focused mainly on the marginal analysis of each trait (with possible covariates). Sometimes, a single-nucleotide polymorphism (SNP) might influence a trait through some intermediate traits, which means the effect on the trait is indirect. It is well-known that marginal analysis cannot distinguish direct and indirect effects. In contrast, conditional analysis can, and thus might be more useful. For example, although low-density lipoprotein (LDL) cholesterol remains the primary treatment target to reduce cardiovascular disease (CVD) risk, it has been shown that high triglyceride (TG) level is a possible independent risk factor for CVD, even in patients with treatment-controlled LDL levels (Boekholdt *et al.* 2012; Sampson *et al.* 2012; Do *et al.* 2013; Toth 2016). Hence, to identify new targets for new therapy to reduce levels of TG or other non-LDL lipids, it is more desirable to detect genetic variants associated with TG and non-LDL lipids after adjusting for LDL; otherwise, due to the known correlations among various lipids, an identified association, say with TG, could simply be due to an indirect effect through LDL.

Importantly, since we most often have summary association statistics only from meta- or mega-analyzed large genome-wide association study (GWAS) data, rather than individual level data, it would be useful to have a pleiotropy test applicable to GWAS summary statistics. With these considerations, we first propose extending the likelihood-ratio test (LRT) of Schaid *et al.* (2016) with individual-level data to that with only summary statistics for marginal analysis with possible covariates (*i.e.*, exogenous variables). Next, for either individual-level GWAS data or summary statistics, we develop a new testing procedure for conditional analysis with some traits as both responses and predictors in multiple conditional regression equations; care has to be taken in this extension to avoid biased estimation and inference. To test a composite null hypothesis, we adopt the union-intersection testing strategy as used in Schaid *et al.* (2016). In particular, we consider conditional regression equations with some traits as both responses and predictors, for which, in addition to extending the LRT of Schaid *et al.* (2016), for simplicity and robustness, we propose using the Wald test in generalized estimating equations (GEE) with the working independence model.

We will use simulations to show the feasibility of marginal or conditional analysis with summary statistics as well as the difference

between marginal and conditional analyses. We will also apply the methods to real data to further demonstrate these points.

Methods

Marginal analysis with individual level data

In this section, we give a brief review of the pleiotropy test of Schaid *et al.* (2016). Throughout this paper, a marginal analysis is defined as one in which no proper subset of traits is used as response in some regression models and as covariate in some other regression models. Hence, a marginal analysis allows the adjustment for nontrait covariates like multiple SNPs. Correspondingly, a conditional analysis is based on multiple regression models including some traits as both responses and covariates.

Suppose that the individual level data contain p traits and q SNPs (or other covariates) for n subjects. As mentioned by Schaid *et al.* (2016), the covariates may be either trait-varying or not trait-varying. For simplicity, we just assume the same q SNPs across different traits. Let $\mathbf{Y}_j = (y_{j1}, y_{j2}, \dots, y_{jn})'$ denote the measured j th trait, and $\mathbf{X}_k = (x_{k1}, x_{k2}, \dots, x_{kn})'$ denote the k th SNP for all n subjects. Assume they are all centered at 0.

The regression model can be expressed as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where

$$\mathbf{Y} = (\mathbf{Y}_1' \ \dots \ \mathbf{Y}_p')',$$

$$\mathbf{X} = (\mathbf{X}_1^* \ \dots \ \mathbf{X}_q^*),$$

$$\mathbf{X}_k^* = \mathbf{I}_p \otimes \mathbf{X}_k,$$

$$\boldsymbol{\beta} = (\boldsymbol{\beta}_1' \ \dots \ \boldsymbol{\beta}_q')',$$

$$\boldsymbol{\beta}_k = (\beta_{k1} \ \dots \ \beta_{kp})',$$

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{\Omega}), \text{ and}$$

$$\boldsymbol{\Omega} = \boldsymbol{\Sigma} \otimes \mathbf{I}_n.$$

\mathbf{I}_n is a $n \times n$ identity matrix, \otimes is the Kronecker product, and the $p \times p$ matrix $\boldsymbol{\Sigma}$ is the covariance matrix for the errors within subjects. All the q SNPs are adjusted for in this model.

Suppose the null hypothesis is H_0 : at most one of the parameters $\beta_{11}, \dots, \beta_{1p}$ is nonzero. It is equivalent to test whether one of the following $p + 1$ tests holds:

$$H_{j0} : \beta_{1j} \neq 0, \beta_{1j_1} = 0 (j_1 \neq j) \text{ and}$$

$$H_{00} : \beta_{1j} = 0 (j = 1, 2, \dots, p)$$

Let $\mathbf{V}_0 = (\mathbf{I}_p \ 0)$ be a $p \times (pq)$ matrix. Denote the matrix obtained after deleting the j th row of \mathbf{V}_0 by \mathbf{V}_j . Then, H_{j0} is

Table 1 Type I errors for set-up A with $\beta_{11} = 1, \beta_{1j} = 0$ ($j \neq 1$)

Sample size	# Traits	ρ	Nominal type-I ER = 0.05				Nominal type-I ER = 0.01			
			LRT-Ind	LRT-Sum	Wald-Ind	Wald-Sum	LRT-Ind	LRT-Sum	Wald-Ind	Wald-Sum
500	4	0.2	0.049	0.048	0.049	0.048	0.013	0.013	0.013	0.013
		0.5	0.044	0.046	0.043	0.046	0.014	0.015	0.014	0.015
		0.8	0.051	0.048	0.050	0.048	0.015	0.015	0.015	0.015
		0.2	0.056	0.054	0.053	0.054	0.013	0.012	0.012	0.012
	10	0.5	0.056	0.052	0.051	0.052	0.013	0.014	0.013	0.014
		0.8	0.055	0.053	0.052	0.053	0.013	0.015	0.013	0.015
		0.2	0.039	0.040	0.039	0.040	0.009	0.008	0.009	0.008
		0.5	0.039	0.041	0.039	0.041	0.010	0.009	0.010	0.009
1000	4	0.8	0.042	0.046	0.042	0.046	0.009	0.008	0.009	0.008
		0.2	0.056	0.060	0.055	0.060	0.016	0.013	0.013	0.013
		0.5	0.056	0.059	0.056	0.059	0.016	0.013	0.014	0.013
		0.8	0.056	0.053	0.054	0.053	0.016	0.012	0.015	0.012
	10	0.2	0.056	0.060	0.055	0.060	0.016	0.013	0.013	0.013
		0.5	0.056	0.059	0.056	0.059	0.016	0.013	0.014	0.013
		0.8	0.056	0.053	0.054	0.053	0.016	0.012	0.015	0.012
		0.2	0.056	0.060	0.055	0.060	0.016	0.013	0.013	0.013

equivalent to $\mathbf{V}_j \boldsymbol{\beta} = 0$. Schaid *et al.* (2016) proposed using the test statistic

$$t_j = \tilde{\mathbf{Y}}' \tilde{\mathbf{X}} (\tilde{\mathbf{X}}' \tilde{\mathbf{X}})^{-1} \mathbf{V}_j' [\mathbf{V}_j (\tilde{\mathbf{X}}' \tilde{\mathbf{X}})^{-1} \mathbf{V}_j']^{-1} \mathbf{V}_j (\tilde{\mathbf{X}}' \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}' \tilde{\mathbf{Y}},$$

where $\tilde{\mathbf{X}} = (\boldsymbol{\Omega}^{-1/2} \mathbf{X}_1^* \dots \boldsymbol{\Omega}^{-1/2} \mathbf{X}_q^*)$, $\tilde{\mathbf{Y}} = \boldsymbol{\Omega}^{-1/2} \mathbf{Y}$, and $\boldsymbol{\Omega} = \boldsymbol{\Sigma} \otimes \mathbf{I}_n$. Use ordinary least squares to estimate $\boldsymbol{\beta}$, and then $\boldsymbol{\Sigma}$ can be estimated using the residuals.

According to Schaid *et al.* (2016), $T = \min_{j=0, \dots, p} t_j$ is the LRT statistic for H_0 .

Since the null distribution of T is complicated, they tested H_0 in two stages. The first stage uses t_0 , and the second stage uses

$$T_1 = \min_{j=1, \dots, p} t_j.$$

t_0 asymptotically follows χ_p^2 when $\boldsymbol{\beta}_1 = 0$. T_1 asymptotically follows χ_{p-1}^2 when only one of $\beta_{11}, \dots, \beta_{1p}$ is nonzero. Reject H_0 only if $t_0 > \chi_p^2(\alpha)$ and $T_1 > \chi_{p-1}^2(\alpha)$.

Marginal analysis with summary statistics

In this section, we describe a few steps to calculate the previous test statistics based only on GWAS summary statistics.

1. If we have each $\hat{\beta}_{kj}$, the marginal effect of \mathbf{X}_k on \mathbf{Y}_j , and its variance $\widehat{\text{var}}(\hat{\beta}_{kj})$, as well as some reference panel and summary statistics of null SNPs (*i.e.*, not associated with any traits), we can estimate $\mathbf{X}_k' \mathbf{Y}_j$, $\mathbf{X}_{k_1}' \mathbf{X}_{k_2}$ and $\mathbf{Y}_{j_1}' \mathbf{Y}_{j_2}$. If we only have Z-statistics, we can regard them as $\hat{\beta}_{kj}$'s and assume $\widehat{\text{var}}(\hat{\beta}_{kj}) = 1$. This approach turns out not to influence the result for marginal analysis, but may change the result for conditional analysis.

1a. Suppose we have individual level data $\mathbf{X}^{\text{ref}} = (\mathbf{X}_1^{\text{ref}} \dots \mathbf{X}_q^{\text{ref}})$ centered at 0 from some reference panel for the SNPs of interest, where $\mathbf{X}_i^{\text{ref}}$ is the i th SNP for n^{ref} subjects. We can estimate $(\mathbf{X}_{k_1}' \mathbf{X}_{k_2})_{q \times q} / n$ by $(\mathbf{X}_{k_1}^{\text{ref}}' \mathbf{X}_{k_2}^{\text{ref}})_{q \times q} / n^{\text{ref}}$, and $(\mathbf{X}_{k_1}' \mathbf{X}_{k_2})_{q \times q}$ by $(\mathbf{X}_{k_1}^{\text{ref}}' \mathbf{X}_{k_2}^{\text{ref}})_{q \times q} n / n^{\text{ref}}$. Note that multiplying $(\mathbf{X}_{k_1}' \mathbf{X}_{k_2})_{q \times q}$ by a constant does not

affect the test results, so we do not really have to use the scalar n/n^{ref} . For marginal analysis with only one SNP, we can simply assume $\mathbf{X}_1' \mathbf{X}_1$ is 1.

1b. Estimate $\mathbf{Y}_j' \mathbf{Y}_j$ using

$$\mathbf{Y}_j' \mathbf{Y}_j - \mathbf{X}_1' \mathbf{X}_1 \hat{\beta}_{1j}^2 = (n-1) \mathbf{X}_1' \mathbf{X}_1 \widehat{\text{var}}(\hat{\beta}_{1j}),$$

where we plug in our estimate of $\mathbf{X}_1' \mathbf{X}_1$ as $\mathbf{X}_1' \mathbf{X}_1$.

1c. Suppose we have Z-statistics \mathbf{Z}_j for marginal analysis of \mathbf{Y}_j with many null SNPs. We can estimate $\mathbf{Y}_{j_1}' \mathbf{Y}_{j_2}$ as $\sqrt{\mathbf{Y}_{j_1}' \mathbf{Y}_{j_1} \mathbf{Y}_{j_2}' \mathbf{Y}_{j_2}} \text{cor}(\mathbf{Z}_{j_1}, \mathbf{Z}_{j_2})$ when $j_1 \neq j_2$, using the idea of Kim *et al.* (2015) and Kwak and Pan (2016, 2017) that

$$\frac{\mathbf{Y}_{j_1}' \mathbf{Y}_{j_2}}{\sqrt{\mathbf{Y}_{j_1}' \mathbf{Y}_{j_1} \mathbf{Y}_{j_2}' \mathbf{Y}_{j_2}}} \approx \text{cor}(\mathbf{Y}_{j_1}, \mathbf{Y}_{j_2}) \approx \text{cor}(\mathbf{Z}_{j_1}, \mathbf{Z}_{j_2}).$$

1d. Estimate $\mathbf{X}_k' \mathbf{Y}_j$ using

$$\mathbf{X}_k' \mathbf{Y}_j = \hat{\beta}_{kj} \mathbf{X}_k' \mathbf{X}_k.$$

2. Estimate the ordinary least squares (OLS) estimate $\hat{\mathbf{B}} = (\hat{\boldsymbol{\beta}}_1 \dots \hat{\boldsymbol{\beta}}_q)$. Note that the OLS estimate and the GLS estimate are the same when there is no trait-specific covariate, which is true in our current scenario. The formula is simply

$$\text{vec}(\hat{\mathbf{B}}) = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{Y}.$$

$\mathbf{X}' \mathbf{Y}$ and $\mathbf{X}' \mathbf{X}$ can be obtained by

$$\mathbf{X}' \mathbf{Y} = (\mathbf{X}_1' \mathbf{Y}_1 \dots \mathbf{X}_1' \mathbf{Y}_p \mathbf{X}_2' \mathbf{Y}_1 \dots \mathbf{X}_2' \mathbf{Y}_p \dots \mathbf{X}_q' \mathbf{Y}_1 \dots \mathbf{X}_q' \mathbf{Y}_p)',$$

$$\mathbf{X}' \mathbf{X} = (\mathbf{X}_{k_1}' \mathbf{X}_{k_2})_{q \times q} \otimes \mathbf{I}_p.$$

3. Estimate $\hat{\boldsymbol{\Sigma}}$ from the residuals. $\hat{\boldsymbol{\Sigma}}$ can be expressed as

Table 2 Type I errors for set-up B with $\beta_{11} = 0, \beta_{1j} = 0 (j \neq 1)$

Sample size	# Traits	ρ	Nominal type-I ER = 0.05				Nominal type-I ER = 0.01			
			LRT-Ind	LRT-Sum	Wald-Ind	Wald-Sum	LRT-Ind	LRT-Sum	Wald-Ind	Wald-Sum
500	4	0.2	0.005	0.005	0.005	0.005	0.001	0.001	0.001	0.001
		0.5	0.005	0.005	0.005	0.005	0.001	0.001	0.001	0.001
		0.8	0.008	0.008	0.008	0.008	0.001	0.001	0.001	0.001
	10	0.2	0.010	0.009	0.008	0.009	0	0	0	0
		0.5	0.010	0.008	0.008	0.008	0	0	0	0
1000	4	0.8	0.008	0.008	0.008	0.008	0	0	0	0.001
		0.2	0.003	0.003	0.003	0.002	0	0	0	0
		0.5	0.004	0.005	0.004	0.005	0.001	0.001	0.001	0.001
	10	0.8	0.006	0.005	0.006	0.005	0.001	0.001	0.001	0.001
		0.2	0.013	0.011	0.012	0.011	0	0	0	0
		0.5	0.011	0.012	0.010	0.012	0	0	0	0
		0.8	0.009	0.009	0.008	0.009	0	0	0	0

$$\hat{\Sigma} = \frac{1}{n-pq} \sum_{i=1}^n (\dot{Y}_i - \hat{B}\dot{X}_i)(\dot{Y}_i - \hat{B}\dot{X}_i)'$$

where $\dot{Y}_i = (y_{1i}, y_{2i}, \dots, y_{pi})'$ and $\dot{X}_i = (x_{1i}, x_{2i}, \dots, x_{qi})'$. By simple algebra, we have

$$(n-pq)\hat{\Sigma} = \sum_{i=1}^n \dot{Y}_i \dot{Y}_i' + \hat{B} \left(\sum_{i=1}^n \dot{X}_i \dot{X}_i' \right) \hat{B}' - \hat{B} \sum_{i=1}^n \dot{X}_i \dot{Y}_i' - \left(\hat{B} \sum_{i=1}^n \dot{X}_i \dot{Y}_i' \right)'$$

$$\sum_{i=1}^n \dot{Y}_i \dot{Y}_i' = (\mathbf{Y}_{j_1}' \mathbf{Y}_{j_2})_{p \times p},$$

$$\sum_{i=1}^n \dot{X}_i \dot{X}_i' = (\mathbf{X}_{k_1}' \mathbf{X}_{k_2})_{q \times q}, \text{ and}$$

$$\sum_{i=1}^n \dot{X}_i \dot{Y}_i' = (\mathbf{X}_k' \mathbf{Y}_j)_{q \times p}.$$

Thus, we can obtain $\hat{\Sigma}$ using $\mathbf{X}_k' \mathbf{Y}_j, \mathbf{X}_{k_1}' \mathbf{X}_{k_2}, \mathbf{Y}_{j_1}' \mathbf{Y}_{j_2}$ and \hat{B} . Then $\hat{\Omega}$ is simply $\hat{\Sigma} \otimes \mathbf{I}_n$.

4. Now we only need to know $\tilde{X}'\tilde{X}$ and $\tilde{X}'\tilde{Y}$. By simple algebra, and replacing Ω with $\hat{\Omega}$, we have

$$\tilde{X}'\tilde{X} = (\mathbf{X}_{k_1}' \mathbf{X}_{k_2})_{q \times q} \otimes \hat{\Sigma}^{-1},$$

$$\tilde{X}'\tilde{Y} = \left((\mathbf{X}_1^{*'} \hat{\Omega}^{-1} \mathbf{Y})' \dots (\mathbf{X}_q^{*'} \hat{\Omega}^{-1} \mathbf{Y})' \right)',$$

$$(\mathbf{X}_k^{*'} \hat{\Omega}^{-1} \mathbf{Y})' = \left(\sum_{j=1}^p a_{1j} \mathbf{X}_k' \mathbf{Y}_j \quad \dots \quad \sum_{j=1}^p a_{pj} \mathbf{X}_k' \mathbf{Y}_j \right),$$

where a_{kj} is the (k, j) th entry of $\hat{\Omega}^{-1} = \hat{\Sigma}^{-1} \otimes \mathbf{I}_n$. Finally, we can calculate the statistics

$$t_j = \tilde{Y}' \tilde{X} (\tilde{X}' \tilde{X})^{-1} \mathbf{V}_j' [\mathbf{V}_j (\tilde{X}' \tilde{X})^{-1} \mathbf{V}_j']^{-1} \mathbf{V}_j (\tilde{X}' \tilde{X})^{-1} \tilde{X}' \tilde{Y}.$$

The remaining testing procedure is the same as that in the previous section.

Conditional analysis with LRTs

For conditional analysis, we consider adjusting for a proper subset of the traits in the regression models for another (non-intersecting) subset of the traits as the responses. The derivation of the above LRT can still be carried out. If we have individual level data, the original method of Schaid *et al.* (2016) should allow incorporation of some traits as trait-specific covariates, though their R package “pleio” allows taking only one SNP without any other covariates. If we only have summary statistics, similar to what was done in Deng and Pan (2017), we can follow the previous idea to derive a similar procedure.

Now, the matrix notation we use is

$$\mathbf{Y} = \mathbf{X}_y \mathbf{G} \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\alpha} \end{pmatrix} + \boldsymbol{\varepsilon},$$

$$\boldsymbol{\varepsilon} \sim N(0, \boldsymbol{\Omega}),$$

$$\boldsymbol{\Omega} = \boldsymbol{\Sigma} \otimes \mathbf{I}_n.$$

\mathbf{Y} is the same as before, while

$$\mathbf{X}_y = \begin{pmatrix} \mathbf{X}_1^* & \dots & \mathbf{X}_q^* & \mathbf{Y}_1^* & \dots & \mathbf{Y}_p^* \end{pmatrix},$$

$$\mathbf{X}_k^* = \mathbf{I}_p \otimes \mathbf{X}_k,$$

$$\mathbf{Y}_j^* = \mathbf{I}_p \otimes \mathbf{Y}_j.$$

Thus, \mathbf{X}_y is a $(np) \times (pq + p^2)$ design matrix combining the information of \mathbf{X} and \mathbf{Y} . $\mathbf{X}_y \mathbf{G}$ is a refined version of \mathbf{X}_y with $(pq + p^2 - p)$ columns.

\mathbf{G} is a $(pq + p^2) \times (pq + p^2 - p)$ matrix obtained by deleting the $(pq + 1)$ th, $(pq + p + 2)$ th, ..., $(pq + p^2)$ th columns of an identity matrix of order $pq + p^2$. The purpose of multiplying

Table 3 Power

# Associated traits	ρ	Nominal type-I ER = 0.05				Nominal type-I ER = 0.01			
		LRT-Ind	LRT-Sum	Wald-Ind	Wald-Sum	LRT-Ind	LRT-Sum	Wald-Ind	Wald-Sum
2	0.2	0.465	0.447	0.441	0.449	0.223	0.203	0.206	0.208
	0.5	0.772	0.752	0.749	0.744	0.516	0.491	0.490	0.499
3	0.2	0.845	0.835	0.829	0.835	0.647	0.618	0.621	0.624
	0.5	0.980	0.979	0.975	0.977	0.927	0.916	0.915	0.914
5	0.2	0.980	0.979	0.979	0.981	0.943	0.933	0.937	0.938
	0.5	1.000	1.000	0.999	0.999	0.994	0.994	0.993	0.990

\mathbf{X}_y by \mathbf{G} is to delete the corresponding columns of \mathbf{X}_y so that the model will be $\mathbf{Y}_j \sim \mathbf{X}_1 + \dots + \mathbf{X}_q + \mathbf{Y}_1 + \dots + \mathbf{Y}_{j-1} + \mathbf{Y}_{j+1} + \dots + \mathbf{Y}_p$ instead of $\mathbf{Y}_j \sim \mathbf{X}_1 + \dots + \mathbf{X}_q + \mathbf{Y}_1 + \dots + \mathbf{Y}_p$.

Denote the effect of the k th SNP on the j th trait by β_{kj} and the l th trait on the j th trait by α_{lj} . $\boldsymbol{\beta} = (\boldsymbol{\beta}_1' \dots \boldsymbol{\beta}_q)'$, $\boldsymbol{\beta}_k = (\beta_{k1} \dots \beta_{kp})'$. Vector $\boldsymbol{\alpha}$ is $p(p-1)$ dimensional. $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1' \dots \boldsymbol{\alpha}_p)'$, $\boldsymbol{\alpha}_l = (\alpha_{l1} \dots \alpha_{l,l-1} \alpha_{l,l+1} \dots \alpha_{lp})'$.

The detailed testing procedure is described below.

Now the test statistic is

$$t_j = \tilde{\mathbf{Y}}' \tilde{\mathbf{X}} (\tilde{\mathbf{X}}' \tilde{\mathbf{X}})^{-1} \mathbf{V}_j' [\mathbf{V}_j (\tilde{\mathbf{X}}' \tilde{\mathbf{X}})^{-1} \mathbf{V}_j']^{-1} \mathbf{V}_j (\tilde{\mathbf{X}}' \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}' \tilde{\mathbf{Y}},$$

where $\tilde{\mathbf{X}} = \boldsymbol{\Omega}^{-1/2} \mathbf{X}_y \mathbf{G}$, $\tilde{\mathbf{Y}} = \boldsymbol{\Omega}^{-1/2} \mathbf{Y}$. Besides, in the new scenario, \mathbf{V}_j is $(p-1) \times (pq + p^2 - p)$ rather than $(p-1) \times (pq)$. We only need to add zeros to the right of the previous \mathbf{V}_j .

Following the previous procedure, we can get t_j 's using a few steps:

1. Estimate $\mathbf{X}_k' \mathbf{Y}_j$, $\mathbf{X}_{k_1}' \mathbf{X}_{k_2}$ and $\mathbf{Y}_{j_1}' \mathbf{Y}_{j_2}$ using the summary statistics, reference panels, and null SNPs. The procedure is described in point (1) in the previous section.
2. Estimate the OLS estimate $\hat{\mathbf{B}}_v$ by

$$\hat{\mathbf{B}}_v = (\mathbf{G}' \mathbf{X}_y' \mathbf{X}_y \mathbf{G})^{-1} \mathbf{G}' \mathbf{X}_y' \mathbf{Y}.$$

Denote $(\mathbf{X}_{k_1}^* \dots \mathbf{X}_{k_q}^*)$ by \mathbf{X}^* and $(\mathbf{Y}_{j_1}^* \dots \mathbf{Y}_{j_p}^*)$ by \mathbf{Y}^* . $\mathbf{X}_y' \mathbf{Y}$ and $\mathbf{X}_y' \mathbf{X}_y$ can be obtained by

$$\mathbf{X}_y' \mathbf{Y} = ((\mathbf{X}^* \mathbf{Y})' \quad (\mathbf{Y}^* \mathbf{Y})')',$$

$$\mathbf{X}_y' \mathbf{X}_y = \begin{pmatrix} \mathbf{X}^* \mathbf{X}^* & \mathbf{X}^* \mathbf{Y}^* \\ \mathbf{Y}^* \mathbf{X}^* & \mathbf{Y}^* \mathbf{Y}^* \end{pmatrix}.$$

Following ideas similar to those presented above, we can easily get

$$\mathbf{X}^* \mathbf{Y} = (\mathbf{X}_1' \mathbf{Y}_1 \dots \mathbf{X}_1' \mathbf{Y}_p \mathbf{X}_2' \mathbf{Y}_1 \dots \mathbf{X}_2' \mathbf{Y}_p \dots \mathbf{X}_q' \mathbf{Y}_1 \dots \mathbf{X}_q' \mathbf{Y}_p)',$$

$$\mathbf{Y}^* \mathbf{Y} = (\mathbf{Y}_1' \mathbf{Y}_1 \dots \mathbf{Y}_1' \mathbf{Y}_p \mathbf{Y}_2' \mathbf{Y}_1 \dots \mathbf{Y}_2' \mathbf{Y}_p \dots \mathbf{Y}_p' \mathbf{Y}_1 \dots \mathbf{Y}_p' \mathbf{Y}_p)',$$

$$\mathbf{X}^* \mathbf{X}^* = (\mathbf{X}_{k_1}' \mathbf{X}_{k_2})_{q \times q} \otimes \mathbf{I}_p,$$

$$\mathbf{X}^* \mathbf{Y}^* = (\mathbf{X}_k' \mathbf{Y}_j)_{q \times p} \otimes \mathbf{I}_p, \text{ and}$$

$$\mathbf{Y}^* \mathbf{Y}^* = (\mathbf{Y}_{j_1}' \mathbf{Y}_{j_2})_{p \times p} \otimes \mathbf{I}_p.$$

3. Estimate $\hat{\boldsymbol{\Sigma}}$ from the residuals. $\hat{\boldsymbol{\Sigma}}$ can be expressed as

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n - p(q + p - 1)} \sum_{i=1}^n (\dot{\mathbf{Y}}_i - \hat{\mathbf{B}} \dot{\mathbf{X}}_i) (\dot{\mathbf{Y}}_i - \hat{\mathbf{B}} \dot{\mathbf{X}}_i)',$$

where $\dot{\mathbf{Y}}_i = (y_{1i}, y_{2i}, \dots, y_{pi})'$ and $\dot{\mathbf{X}}_i = (x_{1i}, \dots, x_{qi}, y_{1i}, \dots, y_{pi})'$. Here, $\dot{\mathbf{X}}_i$ includes both SNPs and traits. Note that $\hat{\mathbf{B}} = (\hat{\boldsymbol{\beta}}_1 \dots \hat{\boldsymbol{\beta}}_q \hat{\boldsymbol{\alpha}}_1^* \dots \hat{\boldsymbol{\alpha}}_p^*)$, while $\hat{\mathbf{B}}_v = (\hat{\boldsymbol{\beta}}_1' \dots \hat{\boldsymbol{\beta}}_q' \hat{\boldsymbol{\alpha}}_1^* \dots \hat{\boldsymbol{\alpha}}_p^*)'$. $\hat{\boldsymbol{\alpha}}_j^*$'s are p dimensional vectors obtained by adding a 0 element between the $(j-1)$ th and j th elements of $\hat{\boldsymbol{\alpha}}_j$'s. Hence, we can get $\hat{\mathbf{B}}$ from $\hat{\mathbf{B}}_v$. By simple algebra, we have

$$[n - p(q + p - 1)] \hat{\boldsymbol{\Sigma}} = \sum_{i=1}^n \dot{\mathbf{Y}}_i \dot{\mathbf{Y}}_i' + \hat{\mathbf{B}} \left(\sum_{i=1}^n \dot{\mathbf{X}}_i \dot{\mathbf{X}}_i' \right) \hat{\mathbf{B}}' - \hat{\mathbf{B}} \sum_{i=1}^n \dot{\mathbf{X}}_i \dot{\mathbf{Y}}_i' - \left(\hat{\mathbf{B}} \sum_{i=1}^n \dot{\mathbf{X}}_i \dot{\mathbf{Y}}_i' \right)',$$

$$\sum_{i=1}^n \dot{\mathbf{Y}}_i \dot{\mathbf{Y}}_i' = (\mathbf{Y}_{j_1}' \mathbf{Y}_{j_2})_{p \times p},$$

$$\sum_{i=1}^n \dot{\mathbf{X}}_i \dot{\mathbf{X}}_i' = \begin{pmatrix} (\mathbf{X}_{k_1}' \mathbf{X}_{k_2})_{q \times q} & (\mathbf{X}_k' \mathbf{Y}_j)_{q \times p} \\ (\mathbf{Y}_j' \mathbf{X}_k)_{p \times q} & (\mathbf{Y}_{j_1}' \mathbf{Y}_{j_2})_{p \times p} \end{pmatrix}, \text{ and}$$

$$\sum_{i=1}^n \dot{\mathbf{X}}_i \dot{\mathbf{Y}}_i' = \left(\left(\mathbf{X}_k' \mathbf{Y}_j \right)_{q \times p'} \quad \left(\mathbf{Y}_{j_1}' \mathbf{Y}_{j_2} \right)_{p \times p} \right)'.$$

Thus we can obtain $\hat{\boldsymbol{\Sigma}}$ using $\mathbf{X}_k' \mathbf{Y}_j$, $\mathbf{X}_{k_1}' \mathbf{X}_{k_2}$, $\mathbf{Y}_{j_1}' \mathbf{Y}_{j_2}$ and $\hat{\mathbf{B}}_v$. Then, $\hat{\boldsymbol{\Omega}}$ is simply $\hat{\boldsymbol{\Sigma}} \otimes \mathbf{I}_n$.

4. Again, we only need to know $\tilde{\mathbf{X}}' \tilde{\mathbf{X}}$ and $\tilde{\mathbf{X}}' \tilde{\mathbf{Y}}$. By simple algebra, and replacing $\boldsymbol{\Omega}$ with $\hat{\boldsymbol{\Omega}}$, we have

$$\tilde{\mathbf{X}}' \tilde{\mathbf{X}} = \mathbf{G}' \left(\left(\begin{pmatrix} (\mathbf{X}_{k_1}' \mathbf{X}_{k_2})_{q \times q} & (\mathbf{X}_k' \mathbf{Y}_j)_{q \times p} \\ (\mathbf{Y}_j' \mathbf{X}_k)_{p \times q} & (\mathbf{Y}_{j_1}' \mathbf{Y}_{j_2})_{p \times p} \end{pmatrix} \otimes \hat{\boldsymbol{\Sigma}}^{-1} \right) \mathbf{G}, \right.$$

$$\left. \mathbf{X}_{k_1}^* \hat{\boldsymbol{\Omega}}^{-1} \mathbf{X}_{k_2}^* = \begin{pmatrix} a_{11} \mathbf{X}_{k_1}' \mathbf{X}_{k_2} & & 0 \\ & \ddots & \\ 0 & & a_{pp} \mathbf{X}_{k_1}' \mathbf{X}_{k_2} \end{pmatrix}, \right.$$

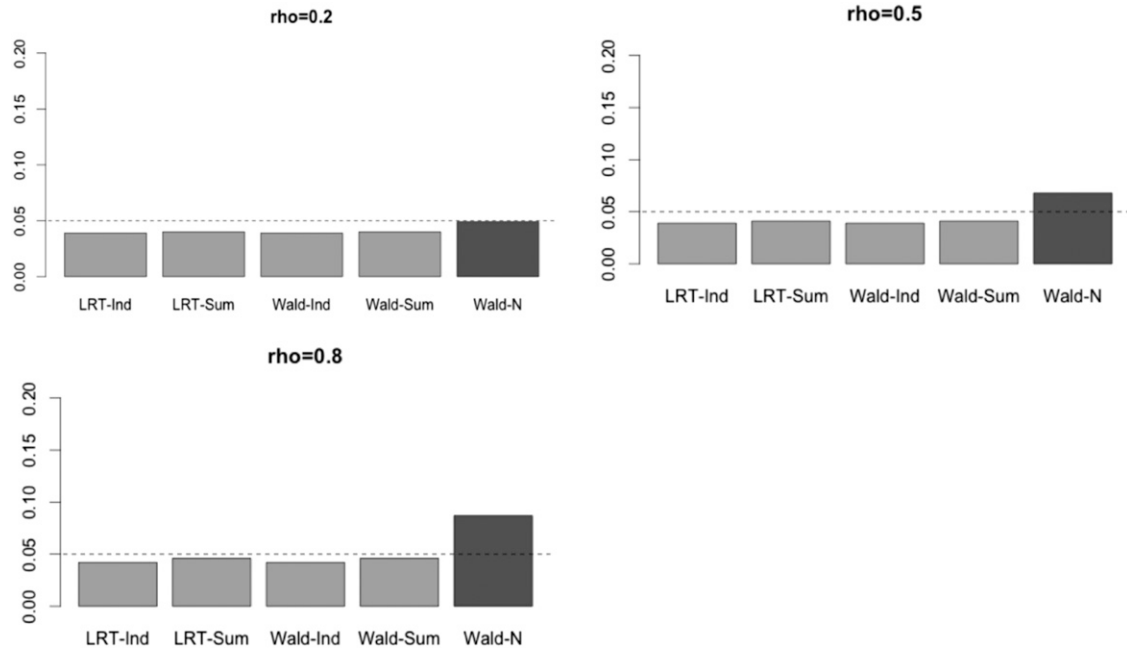


Figure 1 Type I errors in simulation set-up A with $\beta_{11} = 1, \beta_{1j} = 0$ ($j \neq 1$). $n = 1000, p = 4$. Wald-N denotes the GEE Wald test using individual level data and the naive covariance matrix estimate; 1000 null SNPs were generated once for analysis with summary statistics. We used 1000 replications to calculate the rejection rates. A nominal significance level of 0.05 is marked by a dashed line.

$$\mathbf{Y}_{j_1}^{*'} \hat{\boldsymbol{\Omega}}^{-1} \mathbf{Y}_{j_2}^* = \begin{pmatrix} \mathbf{a}_{11} \mathbf{Y}_{j_1}' \mathbf{Y}_{j_2} & & 0 \\ & \ddots & \\ 0 & & \mathbf{a}_{pp} \mathbf{Y}_{j_1}' \mathbf{Y}_{j_2} \end{pmatrix}, \text{ and}$$

$$\mathbf{X}_k^{*'} \hat{\boldsymbol{\Omega}}^{-1} \mathbf{Y}_j^* = \begin{pmatrix} \mathbf{a}_{11} \mathbf{X}_k' \mathbf{Y}_j & & 0 \\ & \ddots & \\ 0 & & \mathbf{a}_{pp} \mathbf{X}_k' \mathbf{Y}_j \end{pmatrix},$$

where a_{lj} is the (l, j) th entry of $\hat{\boldsymbol{\Omega}}^{-1} = \hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{I}_n$. As for $\tilde{\mathbf{X}}' \tilde{\mathbf{Y}}$, we can derive

$$\tilde{\mathbf{X}}' \tilde{\mathbf{Y}} = \mathbf{G}' \left((\mathbf{X}_1^{*'} \hat{\boldsymbol{\Omega}}^{-1} \mathbf{Y})' \dots (\mathbf{X}_q^{*'} \hat{\boldsymbol{\Omega}}^{-1} \mathbf{Y})' (\mathbf{Y}_1^{*'} \hat{\boldsymbol{\Omega}}^{-1} \mathbf{Y})' \dots (\mathbf{Y}_p^{*'} \hat{\boldsymbol{\Omega}}^{-1} \mathbf{Y})' \right),$$

$$(\mathbf{X}_k^{*'} \hat{\boldsymbol{\Omega}}^{-1} \mathbf{Y})' = \left(\sum_{j=1}^p \mathbf{a}_{1j} \mathbf{X}_k' \mathbf{Y}_j \quad \dots \quad \sum_{j=1}^p \mathbf{a}_{pj} \mathbf{X}_k' \mathbf{Y}_j \right),$$

$$(\mathbf{Y}_{j_1}^{*'} \hat{\boldsymbol{\Omega}}^{-1} \mathbf{Y})' = \left(\sum_{j_2=1}^p \mathbf{a}_{1j_2} \mathbf{Y}_{j_1}' \mathbf{Y}_{j_2} \quad \dots \quad \sum_{j_2=1}^p \mathbf{a}_{pj_2} \mathbf{Y}_{j_1}' \mathbf{Y}_{j_2} \right).$$

Hence, we can get everything we need as long as we can estimate $\mathbf{X}_k' \mathbf{Y}_j, \mathbf{X}_{k_1}' \mathbf{X}_{k_2}$ and $\mathbf{Y}_{j_1}' \mathbf{Y}_{j_2}$. These estimates are obtained in step (1).

We can modify the model with a different form of \mathbf{G} . One model we are interested in is $\mathbf{Y}_i \sim \mathbf{X}_1 + \dots + \mathbf{X}_q + \mathbf{Y}_1 + \dots + \mathbf{Y}_{i-1}$. To fit this model, we let \mathbf{G} be a $(pq + p^2) \times (pq + p(p-1)/2)$ matrix. It is obtained by com-

binning the first pq columns of an identity matrix of order $pq + p^2$ with columns $pq + ip + j$, where $i = 1, \dots, p-1, j = 1, \dots, i$. Following this notation, vector $\boldsymbol{\alpha}$ becomes $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_2' \dots \boldsymbol{\alpha}_p)'$, where $\boldsymbol{\alpha}_j = (\alpha_{j1} \dots \alpha_{jj-1})'$.

Our specified model for conditional analysis is a so-called recursive system: (1) it contains a hierarchical model with some traits as predictors for other traits, but not otherwise, and (2) the error terms in the different regression equations are independent; the OLSE is unbiased for a recursive system (Hanushek and Jackson 1977, p. 229).

Conditional analysis with GEE-based Wald tests

The LRT depends critically on the Normality assumption, which may be violated. Instead, we propose using GEE for its much weaker assumptions: for large samples, it depends only on the correct specification of the regression models for the mean function of the traits, but not even on the correct specifications of the variance-covariance matrix of the traits. We propose conducting a union-intersection test with the Wald test in GEE with a working independence model; that is, instead of the GLSE and LRT, we propose using the OLSE and the corresponding sandwich covariance matrix estimate for valid inference.

Using the previous notation, we define the sandwich covariance estimator \mathbf{C} and the naive covariance estimator $\mathbf{C}_{\text{naive}}$ as

$$\mathbf{C} = (\mathbf{X}_y' \mathbf{X}_y)^{-1} (\mathbf{X}_y' \hat{\boldsymbol{\Omega}} \mathbf{X}_y) (\mathbf{X}_y' \mathbf{X}_y)^{-1} \text{ and}$$

$$\mathbf{C}_{\text{naive}} = \hat{\sigma}^2 (\mathbf{X}_y' \mathbf{X}_y)^{-1},$$

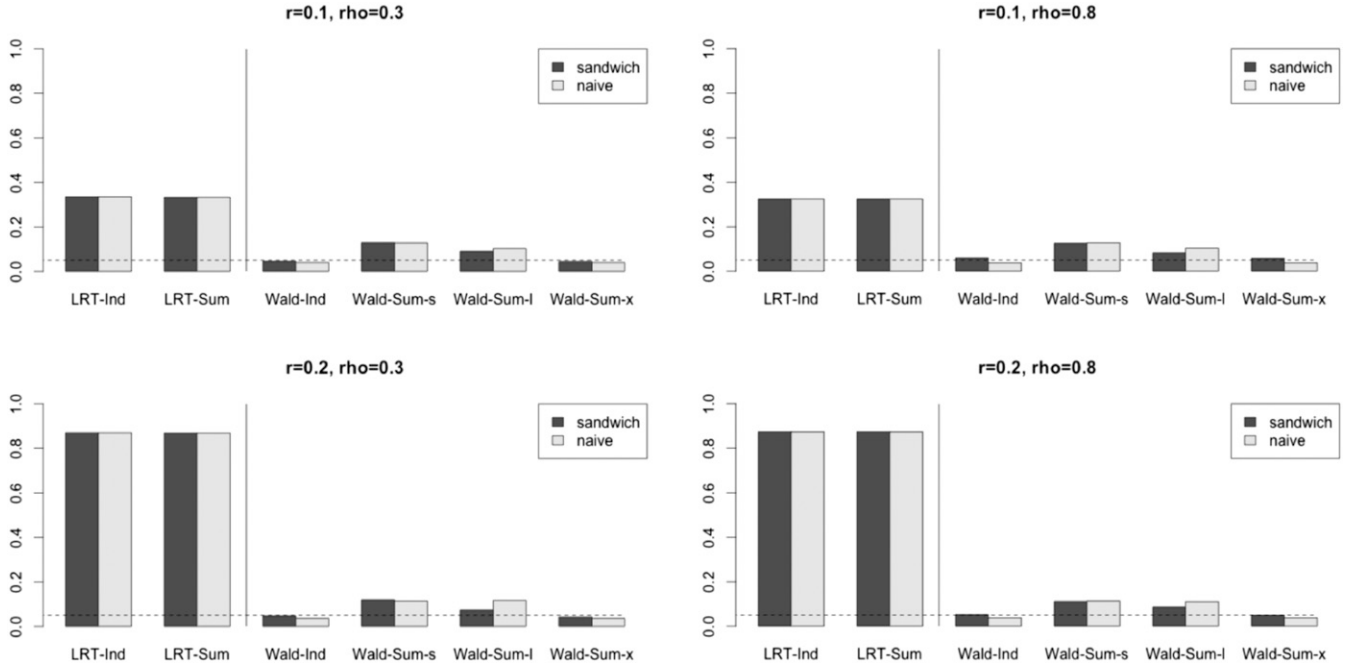


Figure 2 Type I errors with $\beta_{11} = \beta_{22} = 1$, $\beta_{12} = \beta_{21} = 0$, $n = 1000$, and a nominal significance level of 0.05.

where $\hat{\sigma}^2$ is the estimated variance of the error terms for the naive model, which assumes the error terms are identically independently distributed.

The new testing procedure includes a few steps.

1. Estimate $\mathbf{X}_k' \mathbf{Y}_j$, $\mathbf{X}_{k_1}' \mathbf{X}_{k_2}$ and $\mathbf{Y}_{j_1}' \mathbf{Y}_{j_2}$ using the summary statistics, reference panels, and null SNPs.
2. Obtain the OLS estimate $\hat{\mathbf{B}}_v = (\mathbf{G}' \mathbf{X}_y' \mathbf{X}_y \mathbf{G})^{-1} \mathbf{G}' \mathbf{X}_y' \mathbf{Y}$.
3. Estimate $\hat{\Sigma}$ and $\hat{\Omega} = \hat{\Sigma} \otimes \mathbf{I}_n$ from the residuals. The procedure is the same as step 3 in the previous section.
4. Calculate the $(pq + p^2 - p) \times (pq + p^2 - p)$ sandwich covariance matrix estimate

$$\mathbf{C} = (\mathbf{G}' \mathbf{X}_y' \mathbf{X}_y \mathbf{G})^{-1} \mathbf{G}' (\mathbf{X}_y' \hat{\Omega} \mathbf{X}_y) \mathbf{G} (\mathbf{G}' \mathbf{X}_y' \mathbf{X}_y \mathbf{G})^{-1}.$$

where $\mathbf{X}_y' \mathbf{X}_y$ can be obtained in the same way as we did in the previous section, and $\mathbf{X}_y' \hat{\Omega} \mathbf{X}_y$ can be derived by

$$\mathbf{X}_y' \hat{\Omega} \mathbf{X}_y = \begin{pmatrix} (\mathbf{X}_{k_1}' \mathbf{X}_{k_2})_{q \times q} & (\mathbf{X}_k' \mathbf{Y}_j)_{q \times p} \\ (\mathbf{Y}_j' \mathbf{X}_k)_{p \times q} & (\mathbf{Y}_{j_1}' \mathbf{Y}_{j_2})_{p \times p} \end{pmatrix} \otimes \hat{\Sigma}.$$

For the naive method, we can simply use $[n - p(q + p - 1)]S / [2n - p(q + p - 1)]$ to estimate $\hat{\sigma}^2$, where S is the sum of the diagonal elements of $\hat{\Sigma}$. By simple algebra, $[n - p(q + p - 1)]S$ is indeed the sum of squared errors for the naive model.

5. The test statistics are

$$t_j = \hat{\mathbf{B}}_v' \mathbf{V}_j' [\mathbf{V}_j \mathbf{C}^{-1} \mathbf{V}_j']^{-1} \mathbf{V}_j \hat{\mathbf{B}}_v \text{ and}$$

$$T_1 = \min_{j=1, \dots, p} t_j.$$

We reject H_0 only if $t_0 > \chi_p^2(\alpha)$ and $T_1 > \chi_{p-1}^2(\alpha)$. If we want to apply the naive method instead, replace \mathbf{C} with $\mathbf{C}_{\text{naive}} = \hat{\sigma}^2 (\mathbf{X}_y' \mathbf{X}_y)^{-1}$.

Note that this method can also be applied to the analysis adjusting for a few SNPs only, or only for a subset of traits, or their combination. We may extend the methods to sequential tests of multiple associated traits as discussed in Schaid *et al.* (2016) in the future.

Dealing with different samples

Sometimes the summary statistics for different traits may be based on completely different samples (*i.e.*, no overlapping subjects). We can modify the Wald test to carry out the corresponding pleiotropy testing. There are two possible ways.

Approach 1: Note that the OLS estimates for the joint model of all traits vs. SNPs are the same as OLS estimates for separate models of each trait vs. SNPs. Also, since different traits are obtained from different subjects, they are independent, meaning that $\hat{\Sigma}$ should be diagonal. The j th diagonal element of $\hat{\Sigma}$ should simply be the estimated variance of the error term in the regression model for trait j vs. SNPs, which can be calculated by $[se(\hat{\beta}_{1j})]^2 \mathbf{X}_{\text{ind},j}' \mathbf{X}_{\text{ind},j}$, where $\mathbf{X}_{\text{ind},j}' \mathbf{X}_{\text{ind},j}$ is the sum of squares of X for the j th sample (the sample for getting marginal effects on trait j). We can estimate $\mathbf{X}_{\text{ind},j}' \mathbf{X}_{\text{ind},j}$ by cn_i , assuming $\mathbf{X}_{\text{ind},j}' \mathbf{X}_{\text{ind},j}$ is proportional to the sample size.

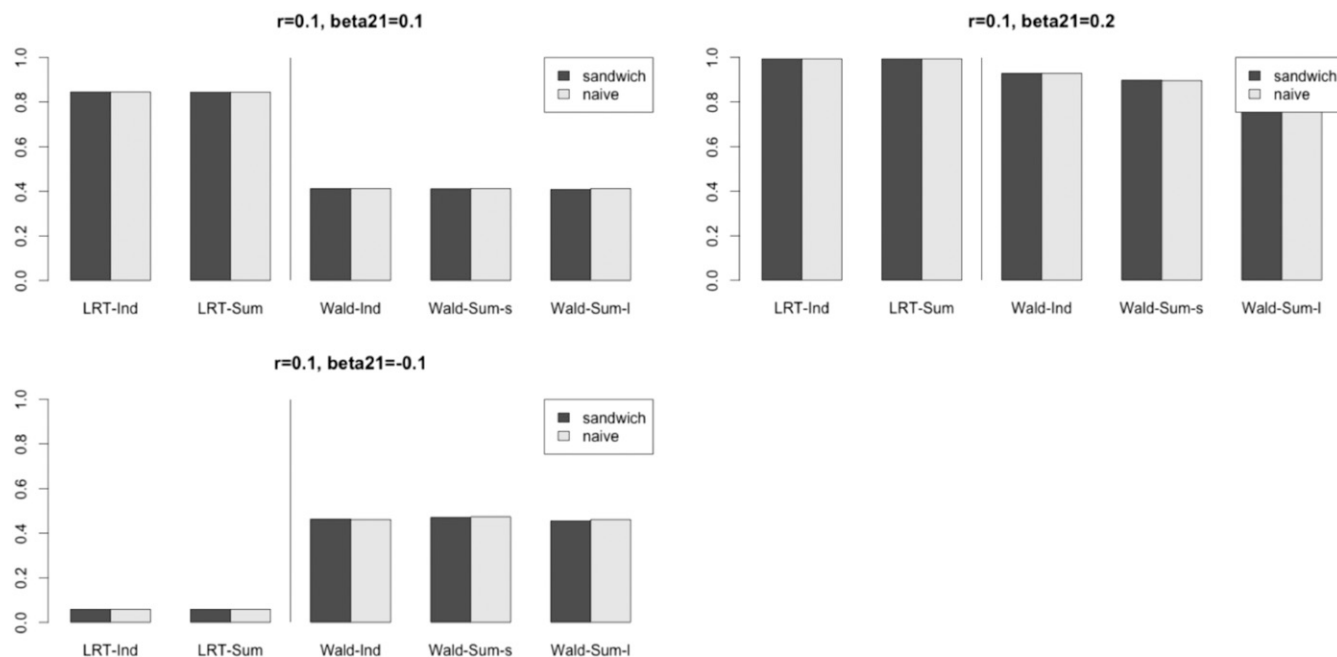


Figure 3 Power with $\beta_{11} = 1, \beta_{22} = 1, \rho = 0.3, \beta_{12} = 0, n = 1000, n^{ref} = 1000$, and a nominal significance level of 0.05.

Since the scale of X does not affect the result, we can simply set c to be 1.

Approach 2: Assume different samples with sample sizes n_j are obtained by a random partition of a whole sample with sample size $n = \sum n_j$. Build the joint model for all $\sum n_j$ subjects so that we do not need to change our previous model. We can assume that $\hat{\beta}_{kj}$ based on all n subjects is the same as that based on just n_j subjects (when the sample size is large). Now $(Y_j'Y_j)^* - (X_1'X_1)^*\hat{\beta}_{1j}^2 = (n_j - 1)(X_1'X_1)^*\widehat{\text{var}}(\hat{\beta}_{1j})$, where $(Y_j'Y_j)^*$ and $(X_1'X_1)^*$ are the $Y_j'Y_j$ and $X_1'X_1$ for n_j subjects. We can also assume $Y_j'Y_j$ for all subjects is approximately $n(Y_j'Y_j)^*/n_j$, $X_1'X_1$ approximately $n(X_1'X_1)^*/n_j$. As a result, we still have $Y_j'Y_j - X_1'X_1\hat{\beta}_{1j}^2 \approx (n - 1)X_1'X_1\widehat{\text{var}}(\hat{\beta}_{1j})$. Hence, we do not need to modify any formulas. We just need to input n as $\sum n_j$.

The two approaches address two different questions. Take a simple example with two traits Y_1, Y_2 and one SNP X . Suppose n_1 subjects were taken to obtain the summary statistics for Y_1 vs. X , and different n_2 subjects were used to calculate the summary statistics for Y_2 vs. X . For the first approach, the

alternative hypothesis is that X is associated with the two populations, from which the n_1 subjects (for Y_1) and the n_2 subjects (for Y_2) were drawn respectively, allowing the two sets of subjects to come from two, possibly different, populations. Rejecting the null hypothesis does not guarantee that X is also associated with the first population's second trait Y_2 . However, for the second approach, the alternative hypothesis is that X influences both Y_1 and Y_2 , assuming all the subjects came from the same population. Hence, the results of the two approaches may be different if the study populations for different traits are indeed different. Nevertheless, in our opinion, the second approach is preferred under the common assumption that the study populations are the same, since it can then tell whether a SNP is indeed associated with more than one trait for the same population, which follows the definition of pleiotropy. Of course, in the second approach we have to assume a common population, which may be violated in practice.

Data availability

The Cotsapas dataset can be downloaded from the online version of Cotsapas *et al.* (2011). The 2010 (Teslovich *et al.*

Table 4 Type I errors with $n = 1000$

		Nominal significance level = 0.05							
		Marginal		Conditional					
β_1	β_3	LRT-Ind	LRT-Sum	LRT-Ind	LRT-Sum	Wald-Ind	Wald-Sum	Wald-Ind-N	Wald-Sum-N
0	0.3	0.042	0.041	0.041	0.041	0.040	0.041	0.040	0.040
	0.6	0.042	0.041	0.041	0.041	0.040	0.041	0.040	0.040
	0.9	0.042	0.041	0.041	0.041	0.040	0.041	0.040	0.040
	1.2	0.042	0.041	0.041	0.041	0.040	0.041	0.040	0.040

10,000 null SNPs were generated once for methods with summary statistics.

Table 5 Power with $n = 1000$

		Nominal significance level = 0.05					
β_1	β_3	Marginal		Conditional			
		LRT-Ind	LRT-Sum	Wald-Ind	Wald-Sum	Wald-Ind-N	Wald-Sum-N
0.1	0.3	0.416	0.416	0.416	0.415	0.414	0.414
	0.6	0.416	0.416	0.416	0.415	0.414	0.414
0.2	0.3	0.955	0.952	0.956	0.955	0.952	0.952
	0.6	0.955	0.953	0.956	0.955	0.952	0.952

2010) and 2013 (Willer *et al.* 2013) lipid data are also publicly available at <http://csg.sph.umich.edu/abecasis/public/lipids2010> and <http://csg.sph.umich.edu/abecasis/public/lipids2013/>, respectively. The R package for our new methods is publicly available at Github: <https://github.com/yangq001/Plei>.

Results

Simulations for marginal analysis

To see whether the marginal analysis using summary statistics can control Type-I errors, we conducted simulations similar to those presented in Schaid *et al.* (2016). In the simulations, only a single SNP with minor allele frequency (MAF) 0.2 was used. The traits were generated from multivariate normal distributions. The variances of the error terms was all 1's with an exchangeable correlation structure with correlation ρ .

Since our method only uses summary statistics, we also generated 1000 null SNPs with the same sample size to estimate the correlation between the traits. We centered the data, and performed a univariate association analysis. Then, we applied our method to the resulting summary statistics. To estimate the rejection rates, we used 1000 replications under each setting.

Now the true model (for set-up A) is

$$\mathbf{Y}_j = \mathbf{X}_1\beta_{1j} + \varepsilon_j \quad (j = 1, \dots, p),$$

$$(\varepsilon_1' \dots \varepsilon_p')' \sim N(0, \mathbf{\Omega}), \text{ and}$$

$$\mathbf{\Omega} = \mathbf{\Sigma} \otimes \mathbf{I}_n,$$

where the diagonal elements of $\mathbf{\Sigma}$ are all 1's and the other elements are ρ . The fitted model for marginal analysis can be simply expressed as $\mathbf{Y} \sim \mathbf{X}_1$.

First, we considered the case when only one trait is associated with the SNP. According to Table 1, all methods had similar performances in terms of type I errors. The results were also close to those in Schaid *et al.* (2016). For convenience, we denote the various analysis methods as:

LRT-Ind: the original LRT of Schaid *et al.* (2016) with individual level data as implemented in R package “pleio.”

LRT-Sum: the LRT with summary statistics.

Wald-Ind: the GEE-based Wald test with individual level data (to calculate $\mathbf{X}\mathbf{X}$ and correlations between traits).

Wald-Sum: the GEE-based Wald test with summary statistics; using a reference panel and some null SNPs to calculate $\mathbf{X}\mathbf{X}$ and correlations between traits.

Next, we considered the case when no trait is associated with the SNP (set-up B). Again, as shown in Table 2, all the methods performed similarly. They seemed to be conservative in this case, which is expected. Recall that the pleiotropy tests are like two-stage tests, and the first stage is the overall test of non-association, which should have a rejection rate of 0.05 (or 0.01) under this setting. The rejection rate of the combined tests is <0.05 (or 0.01), since, to reject pleiotropy, both the overall test and the second-stage test must reject their respective null hypothesis. As a result, all these pleiotropy tests turn out to be conservative, as shown in Schaid *et al.* (2016).

To compare statistical power of the various methods, we let some associated traits have a nonzero effect size of 0.25. The sample size was 500, and the number of traits was 10. As shown in Table 3, the new methods performed similarly to the original methods using individual level data.

Under the first simulation setting, we also carried out the Wald test in GEE (under the working independence model) using the naive covariance estimate, instead of the sandwich covariance estimate. As shown in Figure 1, as expected, this approach, denoted Wald-N, led to inflated type I errors when the error terms were strongly correlated (as the working independence model did not hold). Hence, we recommend using the sandwich covariance.

The above cases included only one SNP. To be more general, we considered a simple scenario with two SNPs, both with MAF 0.2. The correlation between the SNPs is r . We generated the traits using the true model

$$\begin{pmatrix} Y_{1i} \\ Y_{2i} \end{pmatrix} \Big|_{X_{1i}, X_{2i}} \sim N \left(\begin{pmatrix} \beta_{11} & \beta_{21} \\ \beta_{12} & \beta_{22} \end{pmatrix} \begin{pmatrix} X_{1i} \\ X_{2i} \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right).$$

After centering the simulated data, we obtained the summary statistics and conducted analysis adjusting for SNPs. For comparison, we also did an “unadjusted” marginal analysis with only SNP 1. We aimed to test for pleiotropy for SNP 1.

For the methods using only summary statistics, to estimate the correlation of the traits, we generated 1000 null SNPs and obtained their Z-scores by regressing each trait on each null SNP. In addition, we generated SNPs for n^{ref} subjects using MAF 0.2 and correlation r . These were used to estimate $\mathbf{X}_{k_1}'\mathbf{X}_{k_2}$. For each replication, a new reference dataset was generated.

Table 6 Type I errors with $n = 1000, \beta_3 = 0$

β_1	β_2	Nominal significance level = 0.05			
		LRT-Ind	LRT-Sum	Wald-Ind	Wald-Sum
0.3	0.3	0.335	0.335	0.047	0.048
	0.6	0.808	0.807	0.047	0.050
0.6	0.3	0.882	0.882	0.049	0.059
	0.6	1	1	0.049	0.058

1000 null SNPs were generated once for methods with summary statistics. We used the true correlation matrix to estimate $\mathbf{X}_i, \mathbf{X}_j$, and 3000 replications to calculate the rejection rates.

Now the fitted models are $Y_i = \beta_1 X_{1i} + \varepsilon_{1i}$ for the “unadjusted” marginal analysis, and are $Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_{2i}$ for “adjusted” marginal analysis, where each \mathbf{Y} includes the two traits. We considered three versions of the Wald-Sum test:

Wald-Sum-s: using summary statistics and reference data. $n^{\text{ref}} = 1000$.

Wald-Sum-l: using summary statistics and reference data. $n^{\text{ref}} = 10,000$.

Wald-Sum-x: using summary statistics and true $\mathbf{X}_{k_1}, \mathbf{X}_{k_2}$.

Based on Figure 2, it appears that the unadjusted marginal analyses had the highest type I error rates. This is expected because SNP 1 is associated with trait 1, and trait 1 and trait 2 are correlated. As a result, SNP 1 is also marginally associated with trait 2, leading to that the unadjusted marginal analyses rejects the null hypothesis. Our adjusted marginal analysis using summary statistics and true $\mathbf{X}_{k_1}, \mathbf{X}_{k_2}, \mathbf{Y}'\mathbf{Y}$ could control type I errors, but using estimates from reference data resulted in inflated type I errors. The reason was that the estimated $\mathbf{X}_{k_1}, \mathbf{X}_{k_2}$ and the true $\mathbf{X}_{k_1}, \mathbf{X}_{k_2}$ were different. Note that Wald-Ind and Wald-Sum-x almost had the same results, suggesting that using the null SNPs to estimate $\mathbf{Y}'\mathbf{Y}$ was sufficient.

Next, we also looked at the power of different methods. According to Figure 3, as expected, the adjusted methods were more conservative than the unadjusted methods when β_{12} had the same sign as β_{11} , since the adjusted methods could control the type I errors. When we changed the sign of β_{12} , the power of the unadjusted marginal analysis became very low, while the adjusted marginal analysis could maintain its power. Our explanation for this phenomenon is that, when two SNPs are positively correlated but have effects on the trait in different directions, their marginal effects will be diluted, so the unadjusted marginal analysis is less likely to detect pleiotropy. Note that, using the sandwich covariance and the naive covariance, matrix estimates performed similarly in the above settings, even though the two traits were correlated.

Simulations for conditional analysis

Consider a situation with one SNP and two traits. We generated one SNP with MAF 0.2. The true model is

$$Y_{1i} = \beta_1 X_i + \varepsilon_{1i},$$

$$Y_{2i} = \beta_3 X_i + \varepsilon_{2i}.$$

where $(\varepsilon_{1i}, \varepsilon_{2i})$'s ($k = 1, 2; i = 1, 2, \dots, n$) are independent normal random variables with mean 0 and variance 1 (and correlation 0).

Now the fitted model for marginal analysis is $Y_{1i} = \beta_1 X_i + \varepsilon_{1i}, Y_{2i} = \beta_3 X_i + \varepsilon_{2i}$. For conditional analysis, the model is $Y_{1i} = \beta_1 X_i + \varepsilon_{1i}, Y_{2i} = \beta_3 X_i + \beta_2 Y_{1i} + \varepsilon_{2i}$. We examined the estimated type I errors and power of marginal analysis and conditional analysis. Denote the GEE-based Wald tests by Wald-Ind-N and Wald-Sum-N after replacing the sandwich covariance matrix estimate with the naive covariance matrix estimate.

As shown in Table 4, all the methods could control type I errors. As expected, replacing the sandwich covariance estimate by the naive covariance estimate did not really influence the result; since the error terms were uncorrelated, the working independence model used in GEE held, leading to the good performance of the naive covariance estimate.

As shown in Table 5, all methods gave similar power in this case. We noticed that, when $\beta_3 > \beta_1$, increasing β_3 did not increase the power. One possible explanation is that, in this situation, the pleiotropy tests largely depend on whether β_1 is significant. The overall test and the test for β_3 are almost always significant, so only the test for β_1 influences the combined test. When β_3 is increased, the estimates of β_1 and its SE do not change much for marginal analysis, so the power does not improve. Sometimes, the estimated SE for β_1 may become larger, making the test for β_1 less likely to reject the null. The extent of this is slightly greater for the conditional model. As a result, the conditional methods may even lose a little power.

In another scenario, suppose the SNP has an indirect effect on one trait through another trait. Now the model we use to simulate traits is

$$Y_{1i} = \beta_1 X_i + \varepsilon_{1i} \text{ and}$$

$$Y_{2i} = \beta_2 Y_{1i} + \beta_3 X_i + \varepsilon_{2i}.$$

where $(\varepsilon_{1i}, \varepsilon_{2i})$'s ($k = 1, 2; i = 1, 2, \dots, n$) are bivariate normal random variables with variance 1 and correlation 0. They are independent of the SNP.

Again we centered the data and conducted different ways of analysis. The fitted model for marginal analysis is $Y_{1i} = \beta_1 X_i + \varepsilon_{1i}, Y_{2i} = \beta_3 X_i + \varepsilon_{2i}$. For conditional analysis, the model is $Y_{1i} = \beta_1 X_i + \varepsilon_{1i}, Y_{2i} = \beta_3 X_i + \beta_2 Y_{1i} + \varepsilon_{2i}$.

Table 7 Numbers of significant SNPs (P -value <0.01)

	CPMA	LRT-Sum with Approach 2 (overlap)	Wald-Sum with Approach 1 (overlap)	Wald-Sum with Approach 2 (overlap)
With assumption	47	57 (46)	48 (47)	57 (46)
Without assumption		52 (43)	46 (45)	52 (43)

“With assumption”: assuming each SNP is already known to be significant for one trait, then testing whether there is at least one more significant association; “Without assumption”: directly testing whether there are at least two significant associations for each SNP; “Overlap”: the numbers of the significant SNPs that are also among the 47 detected ones by CPMA (“With assumption”).

As shown in Table 6, the new methods could control type I errors when using individual level data, but might have slightly inflated type I error rates when using summary statistics. In contrast, the marginal analysis could not control type I errors at all. Again, using the naive covariance or the sandwich covariance estimate did not matter in this case.

The Cotsapas data

We looked at the data offered by Cotsapas *et al.* (2011). The dataset contains marginal Z-scores of 107 SNPs on each of seven autoimmune diseases: celiac disease (CeD), Crohn’s disease (CD), multiple sclerosis (MS), psoriasis (Ps), rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), and type 1 diabetes (T1D). The original paper found many SNPs with pleiotropy using their proposed CPMA. For comparison, we applied various pleiotropy tests.

Note that this dataset does not contain SE, so we took the Z-scores as effect sizes and assumed the SE was 1. In addition, the summary statistics for different traits used different samples, so we could apply the two approaches discussed earlier for the Wald method.

We applied different approaches to the Cotsapas data (107 SNPs, seven traits). First, we assumed we already knew that each SNP was significant for one specific trait. Then, we tested whether more traits were associated, and compared the result with CPMA. Then, we tested pleiotropy without that assumption.

As shown in Table 7, our proposed Approach 2 detected more SNPs with pleiotropy, and covered most of the significant SNPs identified by CPMA. In this situation with marginal analysis, the Wald test and the LRT gave very similar results. When we directly applied CPMA “without assumption,” 93 SNPs became significant, which, however, were obtained because, in this case, CPMA was testing whether there was at least one association, not necessarily pleiotropy.

The lipid data

The Global Lipids Genetics Consortium GWAS study (Willer *et al.* 2013) has shown many loci associated with more than one trait among LDL, HDL, TG and TC. To study this further, we conducted pleiotropy tests for single SNPs. We applied the methods to two summary association datasets based on

Table 8 Estimated correlations between the lipid traits

	TG	LDL	HDL	TC
TG		0.228	−0.414	0.324
LDL			−0.087	0.873
HDL				0.138

~100,000 and ~189,000 subjects, respectively (Teslovich *et al.* 2010; Willer *et al.* 2013), which we call the 2010 and 2013 data, respectively. The 2013 data are an expanded version of the 2010 data with more study subjects.

First, following the idea of Kim *et al.* (2015), we used the Z-scores of 2,371,319 nonsignificant SNPs for all traits from the 2013 data to estimate the correlations among the four types of lipids, confirming their moderate to high correlations, which imply possible differences between marginal and conditional analyses. The results are shown in Table 8.

Next we conducted a genome-wide scan on each dataset. We tested pleiotropy for 2,363,472 SNPs that are included in both 2010 Lipids data and 2013 Lipids data. These SNPs are also present in the 1000 Genomes Project data (The 1000 Genomes Project Consortium *et al.* 2015). We applied both marginal analysis and conditional analysis. As some previous studies have shown, LDL-lowering treatments might increase the risk of type 2 diabetes partly due to their ontarget mechanisms (Swerdlow *et al.* 2015; Hemani *et al.* 2016). It might be useful to identify new drug targets by identifying genetic variants with pleiotropic effects on non-LDL lipids. To exclude the indirect effects of a genetic variant through LDL, we chose to conduct a conditional analysis of other lipids after adjusting for LDL. The marginal model is

$$LDL_i = \beta_1 X_i + \varepsilon_{1i},$$

$$TC_i = \beta_2 X_i + \varepsilon_{2i},$$

$$TG_i = \beta_3 X_i + \varepsilon_{3i}, \text{ and}$$

$$HDL_i = \beta_4 X_i + \varepsilon_{4i}.$$

In contrast, the conditional model is

$$LDL_i = \beta_5 X_i + \varepsilon_{5i},$$

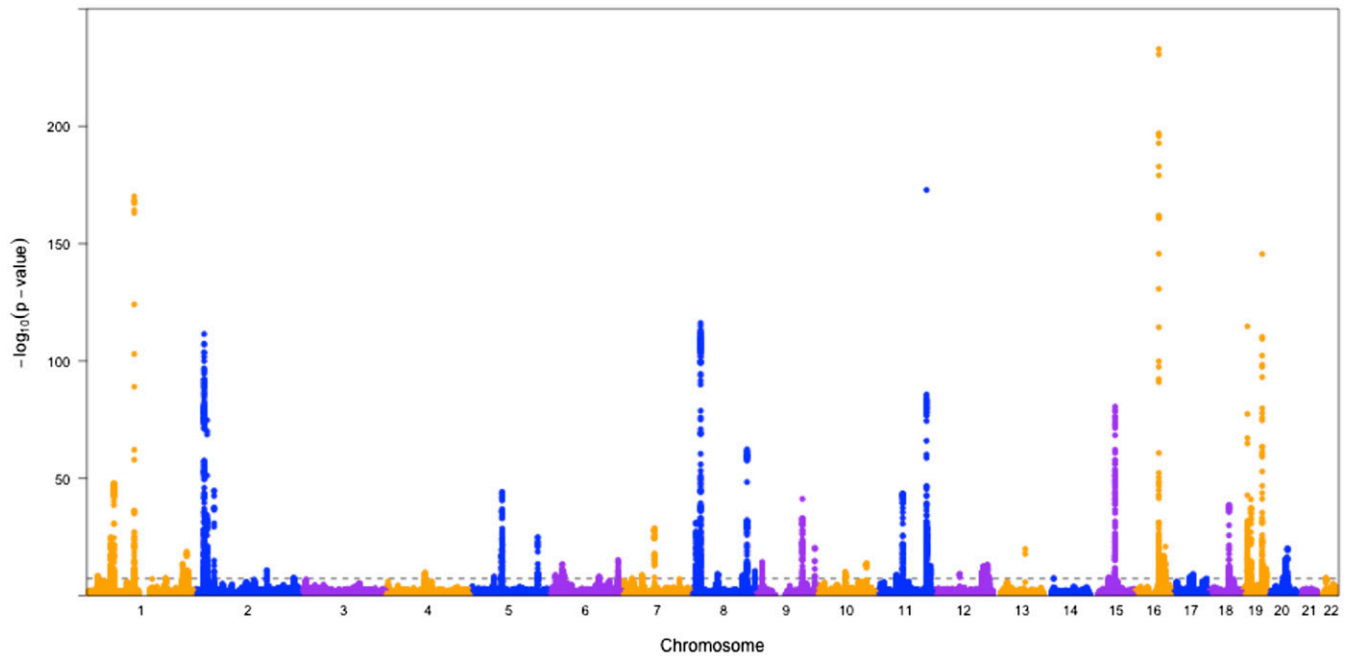
$$TC_i = \beta_6 X_i + \alpha_6 LDL_i + \varepsilon_{6i},$$

$$TG_i = \beta_7 X_i + \alpha_7 LDL_i + \varepsilon_{7i}, \text{ and}$$

$$HDL_i = \beta_8 X_i + \alpha_8 LDL_i + \varepsilon_{8i}.$$

Note that a pleiotropic effect detected by marginal analysis could be due to indirect effect through LDL: for example, if an SNP is causal for high LDL but not for any other three traits, then, by the correlations among the traits, it will be *marginally* associated with all the traits. However, with the conditional model, we can avoid such false positives; we aim to test for

2010 Marginal



2010 Conditional

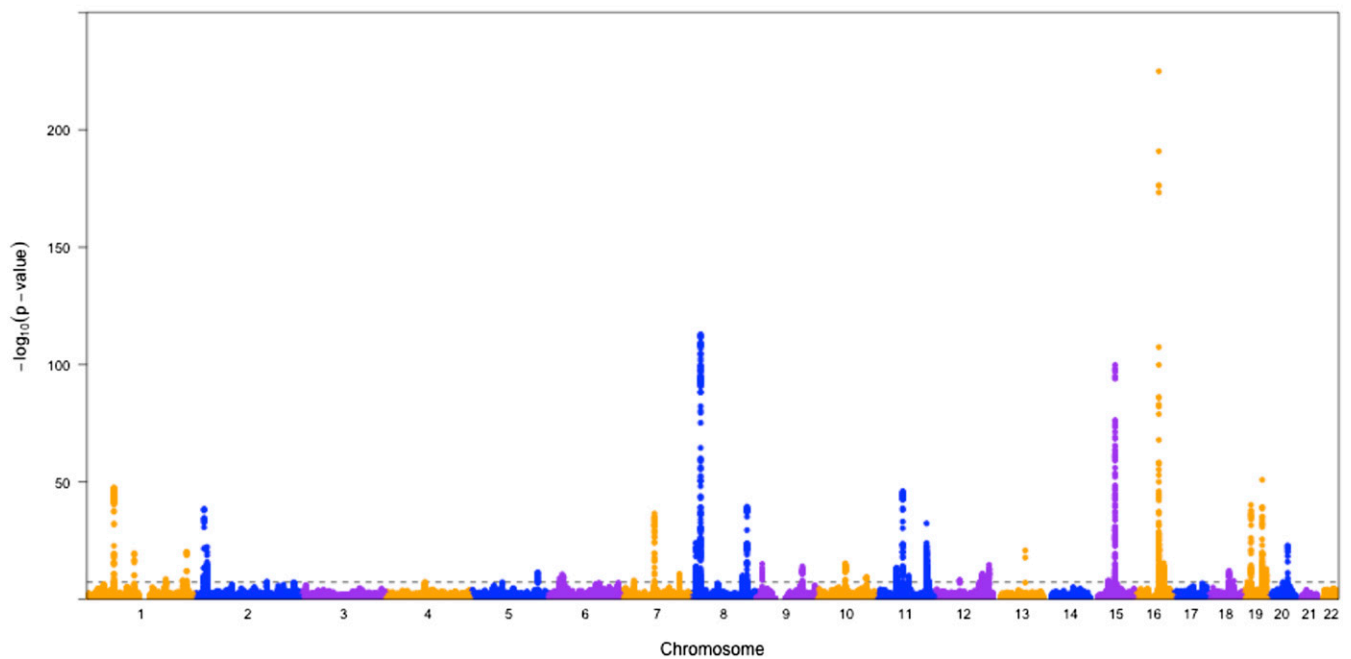


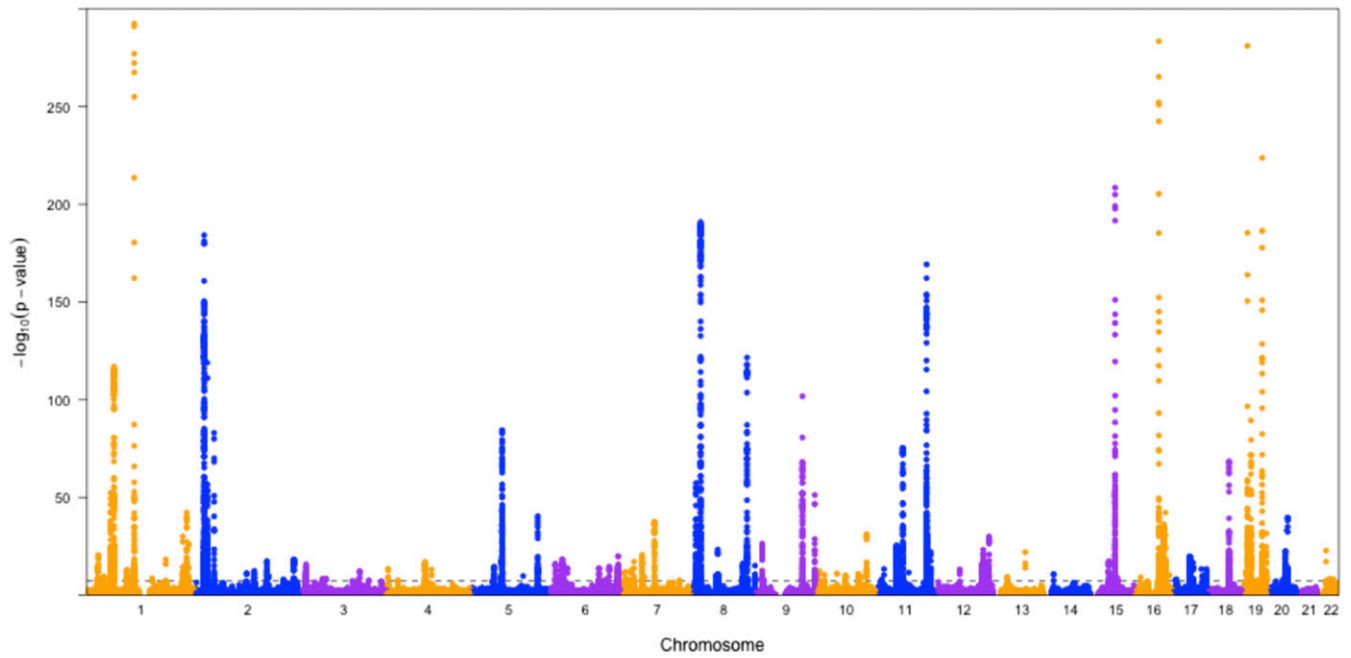
Figure 4 Manhattan plots for pleiotropic testing with the 2010 lipid data.

pleiotropic effects after adjusting for LDL, which may help identify new targets for new drugs as alternatives to LDL-lowering drugs like statin. The null hypothesis for both marginal and conditional pleiotropic tests is that none, or only one, of β_1 , β_2 , β_3 , and β_4 is nonzero.

To show the difference between conditional analysis and marginal analysis, we generated some Manhattan plots. As

shown in Figure 4 and Figure 5, generally the results from 2010 data and 2013 data agreed with each other, except that the latter results seemed more statistically significant, which is expected given the larger sample size of the 2013 data. Interestingly, most SNPs that were highly significant in marginal analysis became less significant in conditional analysis, which means that part of their associations might be indirect

2013 Marginal



2013 Conditional

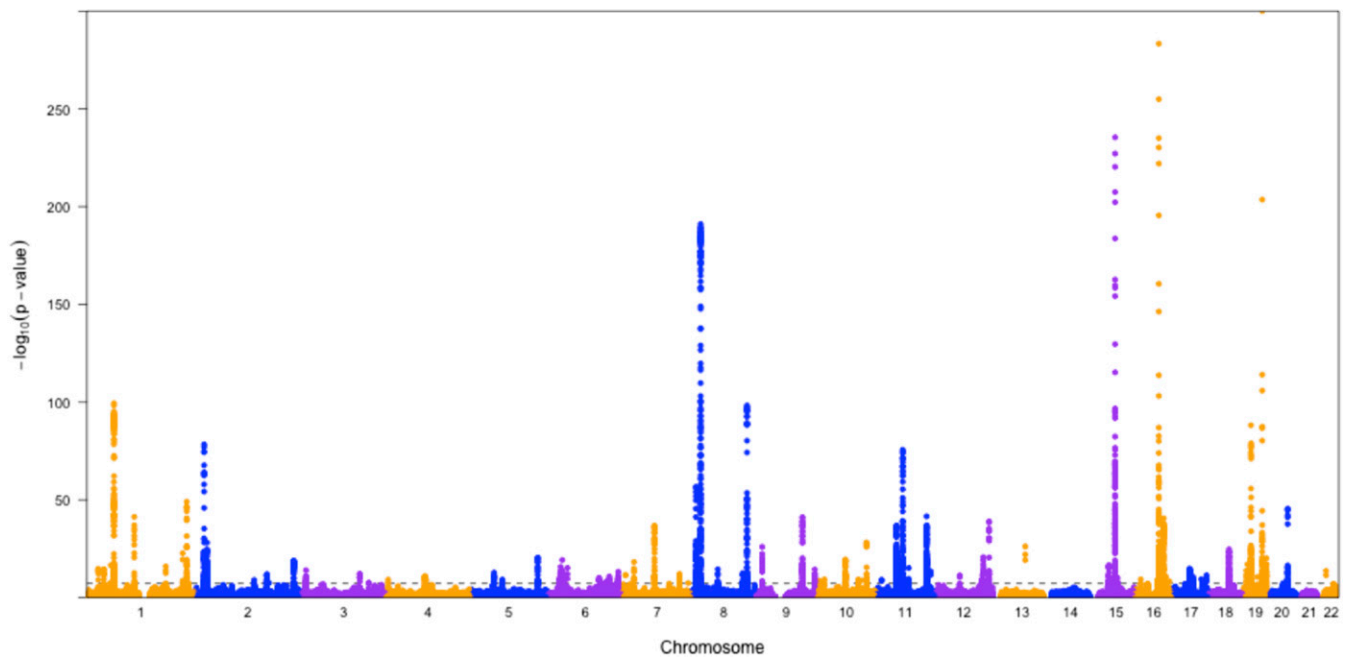


Figure 5 Manhattan plots for pleiotropic testing with the 2013 lipid data.

through LDL as an intermediate. Take SNP rs10043960 on chromosome 5 as an example. The P -value for the pleiotropy test in marginal analysis was 9.2×10^{-10} , while it became nonsignificant at 0.74 after adjusting for LDL.

Note that if the conditional model is true, and $LDL_i = \beta_5 X_i + \varepsilon_{5i}$, $TC_i = 0 \cdot X_i + \alpha_6 LDL_i + \varepsilon_{6i}$, which means the SNP does not affect TC directly, then $E(TC_i | X_i) =$

$E[E(TC_i | LDL_i, X_i) | X_i] = E(\alpha_6 LDL_i | X_i) = \alpha_6 \beta_5 X_i$. As a result, the coefficient β_2 in the marginal model $TC_i = \beta_2 X_i + \varepsilon_{2i}$ should be $\beta_2 = \alpha_6 \beta_5$. Also notice that $Cov(LDL_i, TC_i) = Cov(LDL_i, \alpha_6 LDL_i + \varepsilon_{6i}) = \alpha_6 Var(LDL_i)$, and $Cor(LDL_i, TC_i) = \alpha_6 \sqrt{Var(LDL_i) / Var(TC_i)}$. If we assume $Var(LDL_i) \approx Var(TC_i)$, (which may or may not hold since we do not have individual-level data to verify,) then we have $\alpha_6 \beta_5 \approx Cor(LDL_i, TC_i) \beta_5$,

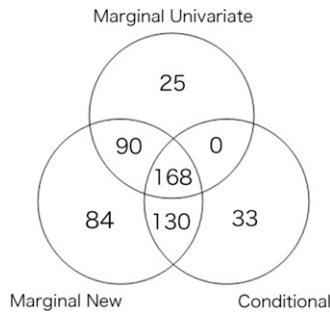


Figure 6 The numbers of the significant loci with pleiotropy detected by the three methods.

suggesting that the estimated marginal effect should be approximately equal to the product of the estimated conditional effect and the correlation between the two traits. For SNP rs10043960, the marginal effect sizes on TG, LDL, HDL, and TC were 0.0096, 0.0402, -0.0050 , and 0.0401, respectively. If we multiply the effect size on LDL by the estimated correlation between LDL and each of the other traits (shown in Table 8), we get 0.0092, -0.0035 , and 0.0351 for TG, HDL, and TC, which are close to the estimated marginal effects. This may suggest that indirect effects through LDL can explain the marginal effects of the SNP on TG, HDL, and TC.

Meanwhile, the significance levels of some SNPs did not change, or even increased after adjusting for LDL. For SNP rs7259679 on chromosome 19, the P -value for marginal pleiotropic effects was 4.0×10^{-6} , while the conditional P -value was 6.3×10^{-13} . The sign of the effect on LDL and the signs of the effects on TG and TC were opposite, while the estimated correlations between LDL, and TG and TC were positive and not small, which resembles the special case discussed above (Table 5). The true effects on LDL, TG, and TC were very likely to be larger than the marginal effects. As a result, the conditional method detected pleiotropy by looking at the true effects, while the marginal method did not.

A simple (and perhaps popular) way to conduct marginal analysis is to look at the marginal P -values for each SNP vs. each trait. If a SNP has at least two significant P -values, we reject the null hypothesis of no pleiotropy. Since this approach involves multiple testing across multiple traits, we apply the Bonferroni adjustment, leading to a genome-wide significance threshold of $5e-8$ divided by the number of traits (*i.e.*, four here). To distinguish this from the marginal analysis described in the previous sections, called “marginal, new,” we call this univariate/single trait testing-based method “marginal, univariate.” Next, we mapped the significant SNPs to loci for each analyses, following the same procedure as used by Schizophrenia Working Group of the Psychiatric Genomics Consortium *et al.* (2014). As shown in Figure 6, most of the loci detected by “marginal univariate” analysis were recovered by the new marginal test, while the latter found many more significant loci. The results of conditional analysis and marginal analysis differed, which might suggest that the effects of certain loci on some traits were indirect, with LDL as an intermediate.

Discussion

We have presented new tests for genetic pleiotropy based on either marginal analysis or conditional analysis using summary statistics. For marginal analysis with only nontrait covariates (*i.e.*, exogenous variables), we extend and apply the LRT of Schaid *et al.* (2016). For conditional analysis with some traits as both responses and predictors (*i.e.*, endogenous variables), in addition to extending the LRT of Schaid *et al.* (2016), for robustness, we also propose using the Wald test in GEE with the working independence model and its corresponding sandwich covariance matrix estimate. Note that we assume a recursive system, for which the OLSE (*i.e.*, GEE estimator with the working independence model) is unbiased (Hanushek and Jackson 1977, p. 229); more general structural equation models (Li *et al.* 2006; P. Wang *et al.* 2016) may require a two-stage least squares estimator, which is much more complex and, more importantly, it is unclear whether it can be applied to summary statistics only. Our extensive simulations showed that our conditional method performed better than marginal analysis in terms of controlling type I errors when indirect effects existed. In some cases, the conditional analysis even had higher power than the marginal analysis. Marginal analysis adjusting for SNPs is also better than unadjusted marginal analysis in certain cases. We applied our methods to the Cotsapas data, and found more SNPs with genetic pleiotropy, covering most of those detected by Cotsapas *et al.* (2011). We also applied different approaches to the 2010 and 2013 lipid data to show possible differences between marginal and conditional analyses.

In the future, we may extend our methods to sequential tests of multiple associated traits as discussed in Schaid *et al.* (2016).

Acknowledgments

We thank the editors and reviewers for many detailed and helpful comments. This research was supported by National Institutes of Health grants R01GM113250, R01HL105397, R01HL116720, and R21AG057038, and by the Minnesota Supercomputing Institute at the University of Minnesota.

Literature Cited

- Berger, R. L., 1997 Likelihood ratio tests and intersection-union tests, pp. 225–237 in *Advances in Statistical Decision Theory and Applications*. Birkhäuser, Boston.
- Boekholdt, S. M., B. J. Arsenault, S. Mora, T. R. Pedersen, J. C. LaRosa *et al.*, 2012 Association of LDL cholesterol, non-HDL cholesterol, and apolipoprotein B levels with risk of cardiovascular events among patients treated with statins: a meta-analysis. *JAMA* 307: 1302–1309.
- Chung, D., C. Yang, C. Li, J. Gelernter, and H. Zhao, 2014 GPA: a statistical approach to prioritizing GWAS results by integrating pleiotropy and annotation. *PLoS Genet.* 10: e1004787.
- Cotsapas, C., B. F. Voight, E. Rossin, K. Lage, B. M. Neale *et al.*, 2011 Pervasive sharing of genetic effects in autoimmune disease. *PLoS Genet.* 7: e1002254.

- Deng, Y., and W. Pan, 2017 Conditional analysis of multiple quantitative traits based on marginal GWAS summary statistics. *Genet. Epidemiol.* 41: 427–436.
- Do, R., C. J. Willer, E. M. Schmidt, S. Sengupta, C. Gao *et al.*, 2013 Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat. Genet.* 45: 1345–1352.
- Hanushek, E. A., and J. E. Jackson, 1977 Statistical methods for social scientists. New York: Academic Press.
- Hemani, G., J. Zheng, K. H. Wade, C. Laurin, B. Elsworth *et al.*, 2016 MR-Base: a platform for systematic causal inference across the phenome using billions of genetic associations. *bioRxiv*. Available at: <https://www.biorxiv.org/content/early/2016/12/16/078972>.
- Kim, J., Y. Bai, and W. Pan, 2015 An adaptive association test for multiple phenotypes with GWAS summary statistics. *Genet. Epidemiol.* 39: 651–663.
- Kim, J., Y. Zhang, and W. Pan Alzheimer's Disease Neuroimaging Initiative, 2016 Powerful and adaptive testing for multi-trait and multi-SNP associations with GWAS and sequencing data. *Genetics* 203: 715–731.
- Kwak, I., and W. Pan, 2016 Adaptive gene- and pathway-trait association testing with GWAS summary statistics. *Bioinformatics* 32: 1178–1184.
- Kwak, I., and W. Pan, 2017 Gene- and pathway-based association tests for multiple traits with GWAS summary statistics. *Bioinformatics* 33: 64–71.
- Li, C., C. Yang, J. Gelernter, and H. Zhao, 2014 Improving genetic risk prediction by leveraging pleiotropy. *Hum. Genet.* 133: 639–650.
- Li, R., S.-W. Tsaih, K. Shockley, I. M. Stylianou, J. Wergedal *et al.*, 2006 Structural model analysis of multiple quantitative traits. *PLoS Genet.* 2: e114.
- Majumdar, A., T. Haldar, and J. S. Witte, 2016 Determining which phenotypes underlie a pleiotropic signal. *Genet. Epidemiol.* 40(5): 366–81.
- Sampson, U. K., S. Fazio, and M. F. Linton, 2012 Residual cardiovascular risk despite optimal LDL cholesterol reduction with statins: the evidence, etiology, and therapeutic challenges. *Curr. Atheroscler. Rep.* 14: 1–10.
- Schaid, D. J., X. Tong, B. Larrabee, R. B. Kennedy, G. A. Poland *et al.*, 2016 Statistical methods for testing genetic pleiotropy. *Genetics* 204: 483–497.
- Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014 Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511: 421–427.
- Solovieff, N., C. Cotsapas, P. H. Lee, S. M. Purcell, and J. W. Smoller, 2013 Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.* 14: 483–495.
- Stephens, M., 2013 A unified framework for association analysis with multiple related phenotypes. *PLoS ONE* 8(7): e65245. Available at: <https://doi.org/10.1371/journal.pone.0065245>.
- Swerdlow, D. I., D. Preiss, K. B. Kuchenbaecker, M. V. Holmes, J. E. Engmann *et al.*, 2015 HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials. *Lancet* 385: 351–361.
- Teslovich, T. M., K. Musunuru, A. V. Smith, A. C. Edmondson, I. M. Stylianou *et al.*, 2010 Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 466: 707–713.
- The 1000 Genomes Project ConsortiumAuton, A., L. D. Brooks, R. M. Durbin, E. P. Garrison *et al.*, 2015 A global reference for human genetic variation. *Nature* 526: 68–74.
- Toth, P. P., 2016 Triglyceride-rich lipoproteins as a causal factor for cardiovascular disease. *Vasc. Health Risk Manag.* 12: 171–183.
- Wang, P., M. Rahman, L. Jin, and M. Xiong, 2016 A new statistical framework for genetic pleiotropic analysis of high dimensional phenotype data. *BMC Genomics* 17: 881.
- Wang, Q., C. Yang, J. Gelernter, and H. Zhao, 2015 Pervasive pleiotropy between psychiatric disorders and immune disorders revealed by integrative analysis of multiple GWAS. *Hum. Genet.* 134: 1195–1209.
- Wang, Y., A. Liu, J. L. Mills, M. Boehnke, A. F. Wilson *et al.*, 2015 Pleiotropy analysis of quantitative traits at gene level by multivariate functional linear models. *Genet. Epidemiol.* 39: 259–275.
- Wang, Z., Q. Sha, and S. Zhang, 2016 Joint analysis of multiple traits using “optimal” maximum heritability test. *PLoS One* 11: e0150975.
- Willer, C. J., E. M. Schmidt, S. Sengupta, G. M. Peloso, S. Gustafsson *et al.*, 2013 Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* 45: 1274–1283.

Communicating editor: E. Eskin