Received: 6 July 2017 | Revised: 27 October 2017 | Accepted: 8 November 2017
DOI: 10.1002/gepi.22105

RESEARCH ARTICLE

WILEY Genetic Epidemiology
OFFICIAL JOURNAL INTERNATIONAL GENETIC EPIDEMIOLOGY SOCIETY www.geneticepi.org

Methods for meta-analysis of multiple traits using GWAS summary statistics

Debashree Ray | Michael Boehnke

Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan, United States of America

Correspondence
Debashree Ray, Department of Epidemiology, Johns Hopkins University, Baltimore, MD, USA.
Email: dray@jhu.edu

Funding information
National Institutes of Health, Grant/Award Numbers: DK062370, HG000376

ABSTRACT

Genome-wide association studies (GWAS) for complex diseases have focused primarily on single-trait analyses for disease status and disease-related quantitative traits. For example, GWAS on risk factors for coronary artery disease analyze genetic associations of plasma lipids such as total cholesterol, LDL-cholesterol, HDL-cholesterol, and triglycerides (TGs) separately. However, traits are often correlated and a joint analysis may yield increased statistical power for association over multiple univariate analyses. Recently several multivariate methods have been proposed that require individual-level data. Here, we develop metaUSAT (where USAT is unified score-based association test), a novel unified association test of a single genetic variant with multiple traits that uses only summary statistics from existing GWAS. Although the existing methods either perform well when most correlated traits are affected by the genetic variant in the same direction or are powerful when only a few of the correlated traits are associated, metaUSAT is designed to be robust to the association structure of correlated traits. metaUSAT does not require individual-level data and can test genetic associations of categorical and/or continuous traits. One can also use metaUSAT to analyze a single trait over multiple studies, appropriately accounting for overlapping samples, if any. metaUSAT provides an approximate asymptotic *P*-value for association and is computationally efficient for implementation at a genome-wide level. Simulation experiments show that metaUSAT maintains proper type-I error at low error levels. It has similar and sometimes greater power to detect association across a wide array of scenarios compared to existing methods, which are usually powerful for some specific association scenarios only. When applied to plasma lipids summary data from the METSIM and the T2D-GENES studies, metaUSAT detected genome-wide significant loci beyond the ones identified by univariate analyses. Evidence from larger studies suggest that the variants additionally detected by our test are, indeed, associated with lipid levels in humans. In summary, metaUSAT can provide novel insights into the genetic architecture of a common disease or traits.

KEYWORDS
cross-phenotype association, GWAS, joint modeling, meta-analysis, METSIM, multiple traits, multivariate analysis, overlapping samples, PheWAS, pleiotropy, score test, summary statistics, T2D-GENES

 © 2017 WILEY PERIODICALS, INC.

1 | INTRODUCTION

Meta-analysis of multiple independent studies is routinely performed to test genetic association of traits by aggregating information on a large number of individuals. Individual data are often not available due to restrictions on data sharing, and hence analysis using summary statistics proves useful. Combining association results from multiple samples of individuals increases statistical power to detect subtle genetic effects. For example, Willer et al. (2013) meta-analyzed lipid traits from 188,577 individuals in 60 studies and detected 62 genome-wide significant loci that were not previously associated with lipid levels in humans.

Although statistical approaches for analysis of individual-level data have moved from the single-trait single-marker paradigm (e.g., Kang et al., 2010) to multiple markers (e.g., Ray, Li, Pan, Pankow, & Basu, 2015; Wu et al., 2011), multiple traits (e.g., Ferreira & Purcell, 2009; Ray, Pankow, & Basu, 2016), and multiple markers and traits (e.g., Basu et al., 2013; Wu & Pankow, 2016), standard approaches for meta-analysis have focused on the analysis of a single trait and a single marker. Many complex disease related traits are correlated. Joint analysis of traits borrows information across all traits and may increase power to detect genetic associations by increasing effective sample size (Diggle, Heagerty, Liang, & Zeger, 2002). For individual-level data, many articles have developed and advocated statistical methods for jointly analyzing correlated traits (see Majumdar, Witte, & Ghosh, 2015; Ray & Basu, 2017; Zhou & Stephens, 2014). Porter and O'Reilly (2017) performed a comprehensive comparison of some of these multitrait methods.

It is only recently that joint meta-analysis of multiple traits using summary statistics has received attention. Stephens (2013) proposed a unified framework for multiple traits single-marker analysis using Bayesian model comparison and model averaging for multivariate regression. This framework allows for approximate testing and explaining genetic associations by using summary statistics. Zhu et al. (2015) proposed a general framework for integrating association evidence using GWAS summary statistics. Their framework can accommodate statistics of multiple continuous or categorical traits, correlated or independent, from a single study or multiple studies. Zhu et al. proposed two tests: S_{Hom} (which assumes equal genetic effect across all traits and studies) and S_{Het} (which allows for trait heterogeneity). Kim, Bai, and Pan (2015) proposed an adaptive sum of powered score (aSPU) test, which lacks a closed form null distribution and depends on Monte Carlo simulations to evaluate P values. Cichonska et al. (2016) proposed metaCCA (where CCA is canonical correlation analysis) that tests association of multiple traits with multiple markers using CCA (Ferreira & Purcell, 2009) framework.

Here, we propose the novel multivariate meta-analysis approach metaUSAT (where USAT is unified score-based association test), a unified score-based association test for the meta-analysis of multiple traits with a single marker using GWAS summary statistics. Current multivariate meta-analysis methods are powerful under certain association patterns (such as sparsity of signals, or homogeneity of signals), and there is a need for a robust association test. metaUSAT is based on the theoretical and empirical findings of Ray et al. (2016) regarding complimentary power performances of CCA/MANOVA (multivariate analysis of variance) and sum of squared score (SSU) tests (Pan, 2009) for individual-level data. Ray et al. (2016) demonstrated that MANOVA may lose significant power when the genetic marker is associated with all the traits, and any test statistic, such as SSU, that does not include the trait correlation structure can be more powerful in such a situation. On the other hand, MANOVA is usually more powerful than other tests when a subset of the correlated traits is associated. The true underlying association scenario (which varies from one genetic marker to another) is not known, and a fixed choice of association test may not be powerful enough. metaUSAT seeks to maximize power by adaptively combining the MANOVA and the SSU tests based solely on the univariate summary statistics. Although both metaMANOVA (the MANOVA test based on summary statistics) and the SSU tests are chi-squared distributed, metaUSAT does not have a closed form null distribution. However, it does not require compute intensive permutations to evaluate P values; instead, we calculate an approximate P-value using a fast one-dimensional numerical integral. metaUSAT retains the flavor of Zhu et al.'s statistics by accommodating summary statistics for continuous and/or binary traits, correlated and/or independent, from one or more studies, which may include overlapping samples. Using metaUSAT, one may perform meta-analysis of a single trait over multiple studies, or multiple traits over one or more studies.

2 | MATERIAL AND METHODS

2.1 | Model and notation

Consider a single GWAS with data on n individuals, genotyped on p genetic variants (say, single nucleotide polymorphisms or SNPs), and measured for K traits. Let $\boldsymbol{Y}_k$ be the $n \times 1$ vector of values for the kth trait and $\boldsymbol{Y}$ be the $n \times K$ matrix of all traits for all individuals. For a given SNP, let $X_i = 0, 1$, or 2 be the number of copies of minor alleles for individual i and $\boldsymbol{X}$ be the $n \times 1$ vector of genotypes for all individuals. For simplicity, we assume there is no other covariate (note that this assumption can be relaxed easily). For the time being, we are interested in testing association between the SNP and the K correlated traits from a single study.

The usual approach is to test for association of each trait separately and report the summary statistics and the P values for each trait based on the marginal/univariate model:

$$\boldsymbol{Y}_k = \boldsymbol{\alpha}_k + \beta_k \boldsymbol{X} + \boldsymbol{\epsilon}_k,\ \ \boldsymbol{\epsilon}_k \sim N_n(\boldsymbol{0}, \sigma_k^2 \boldsymbol{I}_n) \quad \text{for all } k = 1, 2, \ldots, K \tag{1}$$

for continuous traits, or marginal model

$$\text{logit}(P(\boldsymbol{Y}_k = 1|\boldsymbol{X})) = \boldsymbol{\alpha}_k + \beta_k \boldsymbol{X} \ \text{ for all } k = 1, 2, \ldots, K \tag{2}$$

for binary traits. For the kth trait, β_k is the genetic effect and our null hypothesis is $H_{0,k} : \beta_k = 0$. The Wald test statistic for $H_{0,k}$ is $Z_k = \hat{\beta}_k / \text{se}(\hat{\beta}_k)$, where $\hat{\beta}_k$ is the maximum likelihood estimate (MLE) of β_k and $\text{se}(\hat{\beta}_k)$ is its standard error. Under $H_{0,k}$, Z_k has an asymptotic $N(0,1)$ distribution. However, for kth and l th traits, Z_k and Z_l are not independently distributed if the trait correlation is nonzero. In fact, one can show that $\text{corr}(Z_k, Z_l) \approx \text{corr}(Y_k, Y_l)$ when the variability in the estimators of β_k and β_l are ignored (Kim et al., 2015; Zhu et al., 2015).

To test the global null hypothesis of no association with any trait $H_0 : \beta_1 = \ldots = \beta_K = 0$, one can use the summary statistics $\boldsymbol{Z} = (Z_1, \ldots, Z_K)'$. Under H_0, $\boldsymbol{Z}$ has an asymptotic multivariate normal distribution with mean $\boldsymbol{0}$ and covariance matrix $\boldsymbol{R}$, where $\boldsymbol{R}$ is the $K \times K$ correlation matrix of the original traits. Details on estimating $\boldsymbol{R}$ (using single trait summary statistics) are provided in section 2.4.

2.2 | Existing methods

Here, we describe how summary statistics of the K traits for a given SNP can be used to test H_0. Later, in section 2.5, we describe how these methods can be used to conduct meta-analysis using summary statistics from multiple GWAS.

2.2.1 | minP

The minimum P-value (minP) approach selects the most significant result among the K single trait association tests using the test statistic

$$T_{\text{minP}} = \max_{1 \le k \le K} |Z_k|. \tag{3}$$

Its asymptotic P-value accounting for correlated Z statistics (Conneely & Boehnke, 2007) is given by:

$$p_{\text{minP}} = 1 - \text{P}\left(\max\{|Z_1|, ..., |Z_K|\} < t_{\text{minP}}\right) = 1 - \int_{-t_{\text{minP}}}^{t_{\text{minP}}} \cdots \int_{-t_{\text{minP}}}^{t_{\text{minP}}} f_Z(.)\, dz_1 ... dz_K,$$

where $f_Z(.)$ is the multivariate $N_K(\boldsymbol{0}, \hat{\boldsymbol{R}})$ density of $\boldsymbol{Z}$, $\hat{\boldsymbol{R}}$ is the estimate of $\boldsymbol{R}$ and t_{minP} is the observed minP statistic. Computation of p_{minP} requires numerical integration, which can be implemented in `R` using `pmvnorm()` in the `mvtnorm` package (Genz et al., 2016).

2.2.2 | metaMANOVA

An alternative is to carry out a joint analysis of all the Z statistics using a test similar to the multivariate score:

$$T_{\text{metaMANOVA}} = \boldsymbol{Z}' \hat{\boldsymbol{R}}^{-1} \boldsymbol{Z} \underset{H_0}{\overset{a}{\sim}} \chi_K^2. \tag{4}$$

We will call this test metaMANOVA because of its similarity to MANOVA statistic in the context of testing multiple trait association with an SNP using individual-level data (Ray et al., 2016). Although multiple authors (Bolormaa et al., 2014; He et al., 2016; Pausch, Emmerling, Schwarzenbacher, & Fries, 2016) employed this approach, metaMANOVA's type I error and power have not been explored previously for stringent significance levels.

2.2.3 | S_{Hom} and S_{Het}

Zhu et al. (2015) proposed a meta-analysis test S_{Hom} (similar to O'Brien's (1984) test for individual-level data):

$$S_{Hom} = \left(\boldsymbol{1}'(\hat{\boldsymbol{R}}\boldsymbol{W})^{-1}\boldsymbol{Z}\right)' \left(\boldsymbol{1}'(\boldsymbol{W}\hat{\boldsymbol{R}}\boldsymbol{W})^{-1}\boldsymbol{1}\right)^{-1} \left(\boldsymbol{1}'(\hat{\boldsymbol{R}}\boldsymbol{W})^{-1}\boldsymbol{Z}\right) \underset{H_0}{\overset{a}{\sim}} \chi_1^2, \tag{5}$$

where $\boldsymbol{W}$ is a diagonal matrix of weights for the Z statistics, and $\boldsymbol{1}$ is a vector of K ones. Zhu et al. (2015) took sample sizes for the weights. S_{Hom} achieves maximum power when the genetic effects for all traits are equal and in the same direction. Zhu et al. proposed a second statistic S_τ, which seeks to include only Z statistics corresponding to traits with nonzero genetic effects: $S_\tau = (\boldsymbol{1}_\tau'(\hat{\boldsymbol{R}}_\tau \boldsymbol{W}_\tau)^{-1}\boldsymbol{Z}_\tau)'(\boldsymbol{1}_\tau'(\boldsymbol{W}_\tau \hat{\boldsymbol{R}}_\tau \boldsymbol{W}_\tau)^{-1}\boldsymbol{1}_\tau)^{-1}(\boldsymbol{1}_\tau'(\hat{\boldsymbol{R}}_\tau \boldsymbol{W}_\tau)^{-1}\boldsymbol{Z}_\tau)$, where, for a given $\tau > 0$, $\boldsymbol{Z}_\tau$ is the subvector of $\boldsymbol{Z}$ satisfying $|Z_k| > \tau$ and the submatrices $\boldsymbol{W}_\tau$, $\hat{\boldsymbol{R}}_\tau$, $\boldsymbol{1}_\tau$ are defined similarly. For large enough τ, it is possible to have all $|Z_k| < \tau$. In this scenario, set $S_\tau = 0$. Zhu et al. define the test statistic:

$$S_{Het} = \max_{\tau > 0} S_\tau. \tag{6}$$

The null distribution for S_{Het} can be approximated by a gamma distribution and P-value estimated using simulations in Zhu et al.'s `R` program `CPASSOC`.

2.2.4 | aSPU and SSU

Kim et al. (2015) defined the sum of powered score (SPU) test as $SPU(\gamma) = \sum_{k=1}^{K} Z_k^\gamma$, where γ is a positive integer. They constructed multiple $SPU(\gamma)$ tests, with γ values $1, 2, ..., 8$, or ∞, that put more weight on traits with larger Z statistics as γ increases. Kim et al. showed that $SPU(1) = S_{Hom}$. $SPU(2)$, also known as the SSU statistic, is approximately distributed as $a\chi_d^2 + b$ under H_0, where a, b, d can be estimated from

$\hat{\boldsymbol{R}}$ (Pan, 2009). The aSPU test adaptively selects the SPU test with minimum P-value. The SPU(γ) statistics for $\gamma > 2$, and hence the aSPU statistic, do not have closed form null distributions and require Monte Carlo simulations to estimate P values.

2.3 | Proposed method: metaUSAT

In the presence of individual-level data, Ray et al. (2016) proposed a unified score-based association test (USAT) to analyze association of multiple traits with a single SNP. USAT seeks to maximize power by adaptively combining SSU (well suited to scenarios when most or all traits have nonzero genetic affects) and MANOVA (well-suited to most scenarios unless most or all correlated traits are associated). Here, we propose metaUSAT, a meta-analysis version of USAT, that can be calculated using univariate summary statistics. We consider the weighted statistic $T_\omega = \omega T_{\text{metaMANOVA}} + (1-\omega)T_{\text{SSU}},\ \ \omega \in [0,1]$, where $T_{\text{SSU}} = \boldsymbol{Z}'\boldsymbol{Z}$ is the SSU test statistic. Because $T_{\text{metaMANOVA}}$ and T_{SSU} have asymptotic chi-square distributions under H_0 for a given weight ω, T_ω is approximately distributed as a linear combination of (potentially dependent) chi-squared variables. The P-value p_ω of T_ω can be calculated using many algorithms (e.g., Davies, 1980; Liu, Tang, & Zhang, 2009). We define metaUSAT as the weighted combination with the most significant P-value:

$$T_{\text{metaUSAT}} = \min_{\omega \in [0,1]} p_\omega. \tag{7}$$

We consider a grid of 11 equi-spaced values of ω from 0 to 1, and approximate the corresponding P-value using a fast one-dimensional numerical integral (see supplementary S1).

2.4 | Estimation of *R* and its effect on metaUSAT

To estimate the trait correlation matrix $\boldsymbol{R}$, we use the Z-statistics of the SNPs that are not associated with any of the K traits (i.e., SNPs with P values greater than a predefined significance threshold, say 10^{-5}, for any trait). Zhu et al. (2015) showed that under the null hypothesis of no association, the correlation matrix of the univariate summary statistics (obtained by calculating the sample correlation matrix $\hat{\boldsymbol{R}}$ of the $\boldsymbol{Z}'$s over a large number of null SNPs) is the same as the trait correlation matrix. This result holds even in the presence of covariates in (1) or (2) (Liu & Lin, 2017).

It is noteworthy that the performance of metaUSAT and the other afore-mentioned summary statistic based tests depend on the estimation of $\boldsymbol{R}$. In a GWAS, we expect most SNPs to be not associated with any trait, and these null SNPs can be conveniently used to estimate $\boldsymbol{R}$. However, as pointed out by one reviewer, recent evidence from heavily studied complex traits such as height and schizophrenia seems to suggest that these traits are highly polygenic. Consequently, a large portion of the genome in linkage disequilibrium (LD) with the causal variants is also associated with the traits. For the joint analysis of such highly polygenic complex traits using summary statistics, the relation $\text{corr}(Z_k, Z_l) \approx \text{corr}(Y_k, Y_l)$ may not be valid and the estimate of $\boldsymbol{R}$ will be affected. The extent to which this misspecified $\boldsymbol{R}$ affects the validity of the tests depends on the strength of association (of the nonnull SNPs used to estimate $\boldsymbol{R}$) as well as on the structure of the test statistic. Our simulation experiments (see supplementary S4) show that if nonnull SNPs with low to moderate strengths of association are used to estimate $\boldsymbol{R}$, the type I error estimates for metaUSAT and minP are largely unaffected, whereas S_{Hom}, S_{Het}, and metaMANOVA may be heavily affected. It seems to us that test statistics that directly incorporate $\boldsymbol{R}$ (e.g., metaMANOVA) are heavily affected by its misspecification, whereas test statistics incorporating $\boldsymbol{R}$ indirectly only through its null distribution (e.g., minP) are mostly unaffected. The validity of metaUSAT (a data-adaptive minimum P-value approach) is largely unaffected by misspecified estimate of $\boldsymbol{R}$ arising due to polygenicity of traits. It is important to mention that our conclusion is based on a limited simulation experiment. It is beyond the scope of this paper to explore this aspect in more detail.

2.5 | Extension to meta-analysis of multiple GWAS

Consider summary statistics Z_{jk} for association with a given SNP for trait k ($k = 1, 2, \ldots, K$) from study j ($j = 1, 2, \ldots, J$). Some or all J studies may or may not have overlapping samples. Let $\boldsymbol{Z}_j$ be the vector of K summary statistics for study j, $\boldsymbol{Z}$ be the $JK \times 1$ vector of summary statistics from all traits across all studies, and $\boldsymbol{\beta}$ be the corresponding $JK \times 1$ vector of effect sizes. We wish to test $H_0 : \boldsymbol{\beta} = \boldsymbol{0}$ against the two-sided alternative that at least one of the traits has nonzero genetic effect in at least one of the studies.

For kth and lth traits from two studies j and j', Lin & Sullivan (2009) showed that $\text{corr}(Z_{jk}, Z_{j'l}) \approx \frac{n_{jj',kl}}{\sqrt{n_{jk}n_{j'l}}}\text{corr}(Y_{jk}, Y_{j'l})$, where $n_{jj',kl}$ is the number of overlapping samples, in studies j and j', and n_{jk} & $n_{j'l}$ are the sample sizes in the two studies. When the studies are independent ($n_{jj',kl} = 0$), summary statistics from the two studies are uncorrelated. For the perfect overlap scenario ($n_{jj',kl} = n_{jk} = n_{j'l}$), the correlation of summary statistics is approximately same as the correlation of the traits (same as that of a single study with multiple traits). We estimate the $JK \times JK$ correlation matrix $\boldsymbol{R}$ from the JK Z-statistics for the SNPs that do not exceed a predefined significance threshold (say, P-value $= 10^{-5}$) for any trait. The formulation of the $\boldsymbol{Z}$ statistic and the estimation of its correlation in this fashion addresses cryptic relatedness arising from overlapping samples in the studies (Kim et al., 2015; Zhu et al., 2015). Once $\boldsymbol{Z}$ and $\boldsymbol{R}$ are defined, we can use any of the existing methods and metaUSAT.

When meta-analyzing across studies, different studies may have varying sample sizes. Because sample sizes may vary widely across traits and/or studies, we suggest weighting the univariate summary statistics by the corresponding sample sizes. If n_{jk} is the sample size for trait k in study j, we use weighted statistics $\sqrt{n_{jk}}Z_{jk}$ to put more weights on statistics coming from larger studies. Note that this weighting scheme is incorporated in S_{Hom} and S_{Het} statistics in (5) and (6) respectively.

2.6 | Simulation experiments

We conduct simulation experiments to assess type I error and compare power of metaUSAT and the existing methods. For type I error simulations, we consider significance levels $\alpha = 10^{-2}, 10^{-3}, \ldots, 10^{-6}, 5 \times 10^{-7}$. For power simulations, we report empirical powers, based on corrected critical values, at level $\alpha = 10^{-4}$. All analyses used the estimated $\boldsymbol{R}$ based on summary statistics across null replicates.

2.6.1 | Simulation 1: A single study

We generated a single study of $n = 1{,}000$ unrelated individuals, each measured for $K = 5$ or 10 traits and a biallelic SNP X with minor allele frequency (MAF) 0.1 at Hardy-Weinberg equilibrium. For each individual, we simulate K phenotypes using a multivariate normal linear model: $\boldsymbol{Y}_{K\times 1} = \beta_0 \mathbf{1}_{K\times 1} + X\boldsymbol{\beta}_{K\times 1} + \boldsymbol{\epsilon}_{K\times 1}$, where $\beta_0 = 1$ and the error $\boldsymbol{\epsilon}$ is simulated from $N_K(\mathbf{0}, \sigma^2 \boldsymbol{R}(\rho))$. We took $\boldsymbol{R}(\rho)$ as an exchangeable correlation matrix with pair-wise correlation $\rho \in \{0.2, 0.4, 0.6\}$. For type I error simulations, the genetic effects $\boldsymbol{\beta}$ are 0 for all K traits. For power simulations, we choose the genetic effect β_k for an associated trait k so that the SNP explains 0.5% of the trait variance ($k = 1, 2, ..., K$). This, alongwith the MAF of the SNP, determines the genetic effect sizes (see Basu et al., 2013, "Simulations"). We took positive direction of the effect size for all associated traits. The total variance of an associated trait is fixed at 10, which ensures that the variance due to SNP is 0.05 while the residual variance is $\sigma^2 = 9.95$. We wish to test $H_0 : \boldsymbol{\beta} = \mathbf{0}$.

Based on 10^8 null datasets, we estimate type I error of S_{Hom}, S_{Het}, minP, metaMANOVA, and metaUSAT as the proportion of null datasets that give P-value $\leq \alpha$. Our literature search did not yield any article where type I errors of all these summary-statistic-based multivariate methods are studied at a level as low as 5×10^{-7}. We do not consider aSPU for type I error analysis because it requires Monte Carlo simulations, making calculations for 10^8 datasets computationally undesirable. For comparing statistical powers of all methods (including aSPU), we simulate 10^4 nonnull datasets assuming 20–100% of the traits are positively associated with the SNP. To avoid clutter, we are not including SSU (a special case of aSPU) in any of these comparisons.

2.6.2 | Simulation 2: Two independent studies

We consider two independent studies of 1,000 independent individuals, each with measurements on a single SNP with MAF 0.1 and four traits inspired by the METSIM lipids data on total cholesterol (TC), high-density lipoprotein (HDL), low-density lipoprotein (LDL), and triglycerides (TGs). We use the trait correlation matrix $\boldsymbol{R}_{\text{metsim}}$ (supplementary Figure S1(a)) to simulate the four traits using the model described in section 2.6.1 (Simulation 1). We consider five association scenarios: (i) only TC is associated, (ii) TC and LDL are associated, (iii) TC, LDL and TG are associated, (iv) all four traits are associated, and (v) none of the traits is associated. As before, the SNP explains 0.5% of the trait variance when associated. We assume TC, LDL, and TG have negative genetic effects, whereas HDL has positive effect when associated. We simulate two study types: "homogeneous" and "heterogeneous." For "homogeneous" studies, the association pattern of the traits is same across both studies. For "heterogeneous" studies, we assume association scenarios (i)–(iv) in the first study, while the traits are not associated (scenario (v)) in the second study. Supplementary Figure S1(c) shows the estimated correlation matrix. For type I error analysis, we assume scenario (v) for both studies and simulate 10^7 null datasets.

2.6.3 | Simulation 3: Two studies with overlapping samples

We keep everything the same as in Simulation 2 except that the two studies now have 200 overlapping individuals. For "homogeneous" studies, we assume the association pattern is same across the two studies. For "heterogeneous" studies, excluding the overlap, we assume the association scenarios (i)–(iv) in one study, while the traits are not associated (scenario (v)) in the other study. For individuals common to both studies, we assume scenario (v). Supplementary Figure S1(d) shows the estimated correlation matrix, which is similar to the correlation structure of lipid traits from the METSIM and T2D-GENES studies (supplementary Figure S1(b)).

2.7 | Application to lipids data

2.7.1 | METSIM study

The METSIM study is a single-site, longitudinal study of 10,197 men (aged $45-73$) randomly selected from the population of Kuopio, Finland (Stančáková et al., 2009). Participants were genotyped with the Illumina OmniExpress GWAS chip and the Illumina exome chip. Here, we focus on the association statistics of four lipid traits from the first visit: TC, HDL, LDL, TG. Before obtaining the summary statistics, individuals on lipid-lowering medication are removed and TG is log-transformed. The traits are, then, regressed on age and age^2, and residuals are inverse-normalized. We focus on 622,950 autosomal SNPs with MAF$\geq$ 1%. We used kinship matrix in the mixed model framework of EMMAX (Kang et al., 2010) to account for within-ancestry population structure and relatedness.

2.7.2 | T2D-GENES study

The T2D-GENES consortium carried out exome sequencing on 6,504 T2D cases and 6,436 controls from five ancestry groups (Fuchsberger et al., 2016). Here, we consider the 4,541 individuals of European origin, 983 of which are part of the METSIM study sample. As before, we focus on the four lipid traits. Exclusions, transformations, and analysis parallel those for the METSIM lipid traits. Here, we also adjusted sex as a covariate.

3 | RESULTS

3.1 | Simulation 1: A single study

The type I error estimates of metaUSAT and other methods are presented in Table 1. Regardless of the number of traits and the strength of trait correlations, all methods control type I error for moderate levels ($\alpha \geq 10^{-4}$). For more stringent levels, we observe slightly inflated type I errors for all methods except S_{Hom}. The inflation seems to increase with increase in number of traits. We note that type I error of metaUSAT is worst at $\alpha = 5 \times 10^{-7}$; in what follows we correct for this by computing power using empirical threshold. The empirical threshold is based on 10^5 null replicates.

TABLE 1 Simulation 1: Type I error estimates at various significance levels α

		$K = 5$			$K = 10$		
Level α	**Method**	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.6$	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.6$
10^{-2}	S_{Hom}	0.95 [0.95, 0.95]	1.00 [1.00, 1.00]	1.03 [1.03, 1.03]	1.00 [1.00, 1.00]	1.07 [1.07, 1.07]	1.01 [1.01, 1.01]
	S_{Het}	1.05 [1.05, 1.05]	1.05 [1.04, 1.05]	0.98 [0.98, 0.98]	1.01 [1.01, 1.01]	1.06 [1.06, 1.06]	1.02 [1.01, 1.02]
	minP	1.03 [1.03, 1.03]	1.03 [1.03, 1.03]	1.02 [1.02, 1.02]	1.04 [1.04, 1.04]	1.03 [1.03, 1.03]	1.03 [1.03, 1.03]
	metaMANOVA	1.05 [1.05, 1.05]	1.05 [1.04, 1.05]	0.98 [0.98, 0.98]	1.04 [1.04, 1.04]	0.97 [0.97, 0.97]	1.06 [1.06, 1.06]
	metaUSAT	0.82 [0.82, 0.82]	0.92 [0.92, 0.92]	0.93 [0.93, 0.93]	0.93 [0.93, 0.93]	1.01 [1.01, 1.01]	1.06 [1.06, 1.06]
10^{-3}	S_{Hom}	0.93 [0.93, 0.93]	1.00 [1.00, 1.00]	1.03 [1.03, 1.03]	1.00 [1.00, 1.00]	1.11 [1.11, 1.12]	1.02 [1.02, 1.03]
	S_{Het}	1.11 [1.10, 1.11]	1.12 [1.11, 1.12]	1.02 [1.02, 1.02]	1.01 [1.01, 1.02]	1.10 [1.09, 1.10]	1.02 [1.02, 1.03]
	minP	1.05 [1.05, 1.06]	1.05 [1.05, 1.06]	1.05 [1.05, 1.06]	1.07 [1.07, 1.08]	1.06 [1.06, 1.06]	1.07 [1.07, 1.07]
	metaMANOVA	1.09 [1.09, 1.10]	1.08 [1.08, 1.08]	0.99 [0.99, 1.00]	1.08 [1.07, 1.08]	0.98 [0.98, 0.98]	1.11 [1.11, 1.12]
	metaUSAT	0.90 [0.90, 0.90]	1.00 [1.00, 1.00]	1.02 [1.01, 1.02]	1.05 [1.05, 1.05]	1.13 [1.12, 1.13]	1.19 [1.18, 1.19]
10^{-4}	S_{Hom}	0.90 [0.89, 0.91]	1.00 [1.00, 1.00]	1.09 [1.08, 1.10]	1.01 [1.00, 1.02]	1.17 [1.16, 1.18]	1.04 [1.03, 1.05]
	S_{Het}	1.18 [1.17, 1.19]	1.22 [1.21, 1.23]	1.09 [1.08, 1.10]	1.04 [1.03, 1.05]	1.14 [1.13, 1.15]	1.05 [1.04, 1.06]
	minP	1.10 [1.09, 1.11]	1.10 [1.09, 1.11]	1.13 [1.12, 1.14]	1.13 [1.12, 1.14]	1.11 [1.10, 1.12]	1.13 [1.12, 1.14]
	metaMANOVA	1.14 [1.13, 1.15]	1.13 [1.12, 1.14]	1.03 [1.02, 1.04]	1.11 [1.10, 1.12]	0.99 [0.98, 1.00]	1.17 [1.16, 1.18]
	metaUSAT	1.21 [1.20, 1.22]	1.29 [1.28, 1.30]	1.18 [1.17, 1.19]	1.59 [1.58, 1.61]	1.49 [1.48, 1.51]	1.39 [1.38, 1.40]
10^{-5}	S_{Hom}	0.88 [0.85, 0.91]	1.00 [0.97, 1.03]	1.12 [1.09, 1.15]	1.07 [1.03, 1.10]	1.28 [1.25, 1.32]	1.11 [1.08, 1.14]
	S_{Het}	1.31 [1.27, 1.34]	1.36 [1.32, 1.40]	1.17 [1.14, 1.21]	1.12 [1.09, 1.16]	1.27 [1.23, 1.31]	1.12 [1.09, 1.16]
	minP	1.14 [1.10, 1.17]	1.16 [1.12, 1.19]	1.30 [1.27, 1.34]	1.23 [1.20, 1.27]	1.13 [1.09, 1.16]	1.25 [1.21, 1.28]
	metaMANOVA	1.17 [1.14, 1.20]	1.16 [1.12, 1.19]	1.03 [1.00, 1.06]	1.19 [1.16, 1.23]	1.03 [1.00, 1.06]	1.31 [1.27, 1.34]
	metaUSAT	1.38 [1.35, 1.42]	1.45 [1.41, 1.49]	1.28 [1.25, 1.45]	1.92 [1.88, 1.97]	1.74 [1.70, 1.79]	1.58 [1.54, 1.62]
10^{-6}	S_{Hom}	0.80 [0.71, 0.89]	0.92 [0.82, 1.02]	1.11 [1.00, 1.22]	1.12 [1.01, 1.23]	1.49 [1.37, 1.61]	1.26 [1.15, 1.38]
	S_{Het}	1.44 [1.32, 1.56]	1.43 [1.31, 1.55]	1.26 [1.15, 1.37]	1.21 [1.10, 1.32]	1.47 [1.35, 1.59]	1.26 [1.15, 1.38]
	minP	1.31 [1.20, 1.42]	1.27 [1.16, 1.38]	1.64 [1.51, 1.77]	1.30 [1.19, 1.41]	1.16 [1.05, 1.27]	1.45 [1.33, 1.57]
	metaMANOVA	1.19 [1.08, 1.30]	1.10 [0.99, 1.21]	1.00 [0.90, 1.10]	1.32 [1.21, 1.43]	1.23 [1.12, 1.34]	1.54 [1.42, 1.66]
	metaUSAT	1.46 [1.34, 1.58]	1.48 [1.36, 1.60]	1.33 [1.22, 1.45]	2.38 [2.23, 2.53]	2.21 [2.06, 2.36]	2.08 [1.94, 2.22]
5×10^{-7}	S_{Hom}	0.72 [0.60, 0.84]	0.82 [0.69, 0.95]	1.00 [0.86, 1.14]	1.20 [1.05, 1.35]	1.50 [1.33, 1.67]	1.25 [1.09, 1.41]
	S_{Het}	1.30 [1.14, 1.46]	1.42 [1.25, 1.59]	1.32 [1.16, 1.48]	1.30 [1.14, 1.46]	1.38 [1.21, 1.55]	1.17 [1.02, 1.32]
	minP	1.46 [1.29, 1.63]	1.30 [1.14, 1.46]	1.90 [1.71, 2.10]	1.42 [1.25, 1.59]	1.08 [0.93, 1.23]	1.64 [1.46, 1.82]
	metaMANOVA	1.20 [1.05, 1.36]	1.25 [1.09, 1.41]	1.15 [1.00, 1.30]	1.32 [1.16, 1.48]	1.20 [1.05, 1.35]	1.62 [1.44, 1.80]
	metaUSAT	1.54 [1.36, 1.72]	1.74 [1.55, 1.92]	1.54 [1.36, 1.71]	2.52 [2.30, 2.74]	2.42 [2.20, 2.64]	2.18 [1.97, 2.39]

Notes: This table lists the type I error estimates divided by the significance level α and the corresponding $100(1-\alpha)\%$ confidence intervals in brackets. The ideal point estimate for any cell is 1. Estimates are based on 10^8 null datasets, each with K traits, 1 SNP and sample size 1,000.

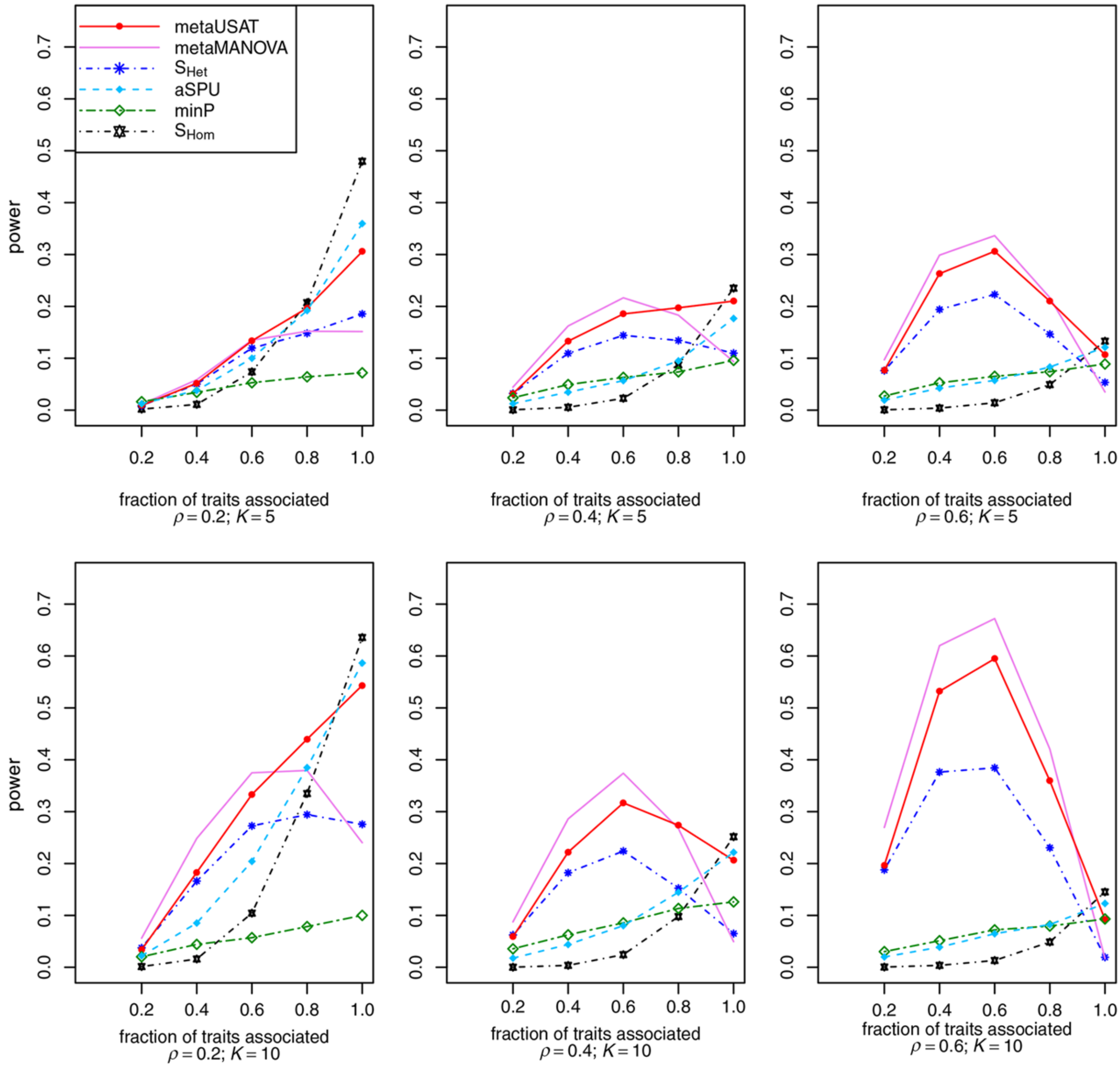


FIGURE 1 Simulation 1: Empirical power curves (based on corrected critical values) of S_{Hom}, S_{Het}, metaMANOVA, metaUSAT, minP, and aSPU at significance level $\alpha = 10^{-4}$

Notes: Power estimates are based on 10^4 datasets with 1,000 unrelated samples. Each sample has $K = 5$ or 10 traits with pairwise trait correlations $\rho = 0.2$, 0.4, or 0.6.

Figure 1 summarizes the empirical powers (based on corrected critical values) of all methods. We observe that as correlation becomes stronger and the number of associated traits increases, S_{Hom}, minP, and aSPU lose power in most association scenarios. S_{Het} is dominated by metaMANOVA, which is usually most powerful. However, metaMANOVA loses power considerably as the proportion of associated traits increases. This phenomenon of metaMANOVA's power loss is the same as what Ray et al. (2016) observed for MANOVA (for analyzing individual-level data) and provided an explanation for. When most or all of the traits are associated, aSPU and S_{Hom} are quite powerful. Irrespective of the number of associated traits and the strength of correlation, metaUSAT, being data-adaptive, has near optimal power to detect association at all association scenarios. Results for genetic marker with MAF 0.5 (not shown) are qualitatively similar. Apart from exchangeable correlation, we also consider an AR1(ρ) correlation structure (autoregressive correlation matrix of order 1 with parameter ρ) and, as before, we find metaUSAT's power to be robust across association scenarios (supplementary Figure S2).

3.2 | Simulation 2: Two independent studies

The estimated correlation matrix, based on 5,000 null summary statistics, is given in supplementary Figure S1(c).

TABLE 2 Simulation 2: Comparison of empirical powers (based on corrected critical values) for two independent studies at level $\alpha = 10^{-4}$

Study type	No. of traits associated	Meta-analysis method					
		S_{Hom}	S_{Het}	minP	aSPU	metaMANOVA	metaUSAT
Homogeneous	1	0.999	0.923	0.034	*0.009*	**1.000**	**1.000**
	2	*0.000*	0.111	0.046	0.082	**0.151**	0.133
	3	0.306	0.254	*0.078*	**0.364**	0.250	0.285
	4	*0.009*	**0.725**	0.100	0.387	0.665	0.632
Heterogeneous	1	0.357	0.661	0.019	*0.004*	**0.995**	0.992
	2	*0.000*	0.016	**0.024**	0.019	0.023	0.020
	3	*0.017*	0.036	0.035	0.044	0.038	**0.045**
	4	*0.001*	**0.194**	0.050	0.068	0.159	0.141

Notes: Power is estimated based on 10^4 nonnull datasets. For a given association scenario, the method with highest power is bold-faced and the method with lowest power is italicized.

TABLE 3 Simulation 3: Comparison of empirical powers (based on corrected critical values) for two studies with overlapping samples at level $\alpha = 10^{-4}$

Study type	No. of traits associated	Meta-analysis method					
		S_{Hom}	S_{Het}	minP	aSPU	metaMANOVA	metaUSAT
Homogeneous	1	0.984	0.852	0.034	*0.018*	**1.000**	**1.000**
	2	*0.000*	0.048	0.051	**0.099**	0.083	0.093
	3	0.244	0.127	*0.077*	**0.300**	0.146	0.216
	4	*0.005*	**0.485**	0.103	0.435	0.456	0.472
Heterogeneous	1	0.101	0.495	0.006	*0.001*	**0.865**	0.807
	2	*0.000*	0.004	**0.009**	0.006	0.007	0.006
	3	*0.007*	0.009	0.011	**0.012**	**0.012**	**0.012**
	4	*0.001*	0.038	0.018	0.026	**0.041**	0.034

Notes: Power is estimated based on 10^4 nonnull datasets. For a given association scenario, the method with highest power is bold-faced and the method with lowest power is italicized.

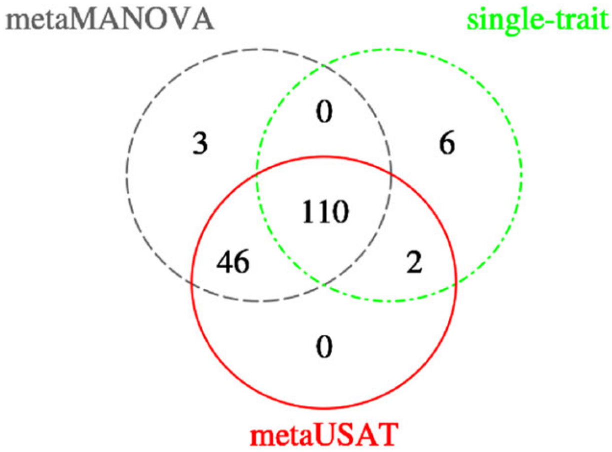
Supplementary Figure S1(a) and (c) show that trait correlations can be approximated by the correlations of summary statistics. Type I error estimates (supplementary Table S1) indicate all methods control type I error for low error levels. Table 2 suggests S_{Het}, metaMANOVA, and metaUSAT are usually most powerful. metaMANOVA and metaUSAT have similar powers. S_{Hom} and minP are least powerful in most cases. aSPU is least powerful when a small proportion of traits is associated. Results for MAF 0.5 (not shown) are qualitatively similar. We also conducted this power comparison for binary traits and found metaUSAT to be robust across association scenarios (supplementary Table S4).

3.3 | Simulation 3: Two studies with overlapping samples

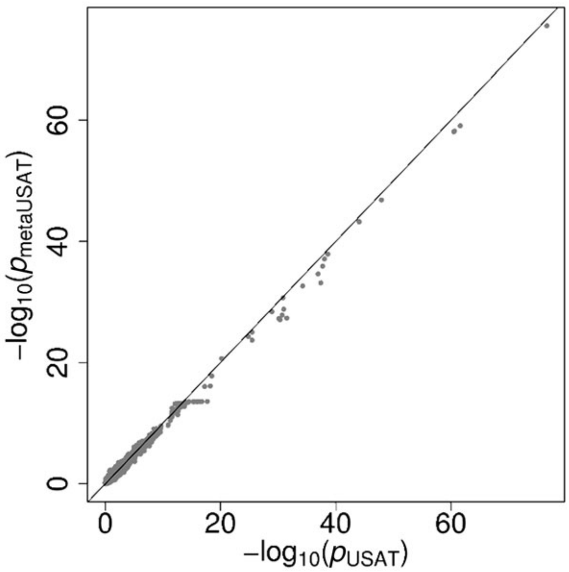
Type I error estimates (supplementary Table S2) are as expected from the earlier type I error analyses. Empirical powers (Table 3) of the methods in the presence of overlapping samples are similar to the simulation without shared individuals (Table 2). We observed similar conclusions when this power comparison is conducted for binary traits (supplementary Table S5).

3.4 | METSIM Study: Joint analysis of lipid traits

Single-trait analysis identified 118 associated variants at the four-trait Bonferroni-corrected threshold of 1.25×10^{-8} (Figure 2(a)). metaMANOVA and metaUSAT, respectively, identified 159 and 158 associated variants at threshold 5×10^{-8}. To identify independent association signals, we grouped significant variants (with pairwise distance < 500 kb) into loci using LD $r^2 > 0.1$. Both metaMANOVA and metaUSAT identified 28 such independent loci, 27 of which (except rs3093032, a 3'-UTR variant in *ICAM1* gene) are known to be associated with lipids from published literature (supplementary Table S6). Additionally, we jointly analyzed individual-level data on these lipid traits using USAT. Figure 2(b) shows concordance of *P* values based on individual-level data and *P* values based on summary statistics.



(a)



(b)

FIGURE 2 METSIM Study: (a) Venn diagram of the number of SNPs (and not independent loci) found significant by each of metaUSAT, metaMANOVA, and single-trait analyses

Notes: A total of 622,950 SNPs (MAF $\geq$ 1%) are tested. For the single-trait analysis, a variant is declared as significant if its *P*-value for at least one trait is $< 1.25 \times 10^{-8}$ (four-trait Bonferroni-corrected GWAS threshold). It should be noted that most of these significant SNPs are in LD. (b) metaUSAT *P* values (joint analysis based on summary data) plotted against USAT *P* values (joint analysis based on individual-level data).

3.5 | METSIM + T2D-GENES studies: Meta-analysis of a single trait from studies with overlapping samples

We tested genetic associations of TC with 31,897 variants (MAF $\geq$ 1%) using summary statistics from METSIM and T2D-GENES studies. metaUSAT, metaMANOVA, and single-trait analyses, respectively, found 12, 12, and 9 SNPs as significant (Figure 3(a)). Published literature indicate that signals identified by metaUSAT (or metaMANOVA) are known to be associated with cholesterol levels (supplementary Table S7). Figure 4(a) plots the metaUSAT *P* values when overlap is present against metaUSAT *P* values when the overlapping individuals are excluded from the METSIM sample. Concordance of the *P* values suggest metaUSAT appropriately accounted for overlapping samples.

3.6 | METSIM + T2D-GENES studies: Joint meta-analysis of lipid traits from studies with overlapping samples

metaUSAT, metaMANOVA, and single-trait analysis, respectively, found 26, 22, and 19 SNPs as significant (Figure 3(b)). metaMANOVA and metaUSAT detected more signals by borrowing information from correlated traits across studies. All of the signals found by both metaMANOVA and metaUSAT are known to be associated with lipid levels in humans from previous studies (supplementary Table S8). All the SNPs detected by metaMANOVA and by independent analysis of each trait were identified by metaUSAT. Further, metaUSAT exclusively reports four significant SNPs (of which three are independent) that metaMANOVA fails to find (Table 4). For these SNPs, we also report the empirical *P* values (calculated using 8.5×10^9 Monte Carlo simulations) to ensure these are not false associations detected as a result of slightly inflated type I error of metaUSAT at stringent error levels. Details of this empirical *P*-value calculation of metaUSAT are provided in supplementary S2. Finally, in Figure 4(b), we again observe concordance of metaUSAT *P* values with and without shared individuals.

4 | DISCUSSION

Most GWAS have focused on testing genetic association to single traits. Several recent articles have advocated the joint analysis of multiple traits for improving statistical power to detect associated genetic variants. In this article, we propose a new method for multivariate meta-analysis, metaUSAT, an extension of our multivariate association test USAT (Ray et al., 2016). For a given genetic variant, metaUSAT tests the association of multiple traits from a single/multiple studies using univariate summary statistics. Importantly, it bypasses the need for individual-level data, which is often unavailable or difficult to obtain.

Our simulation experiments and real data analyses establish that metaUSAT is often more powerful than any of the existing tests for multivariate meta-analysis. It can be especially advantageous in detecting highly pleiotropic variants that simultaneously influence multiple traits. Apart from proposing new method metaUSAT, we also study power and type I error performances of metaMANOVA and other summary statistic based multitrait methods at stringent error levels. metaUSAT and metaMANOVA can accurately control type I

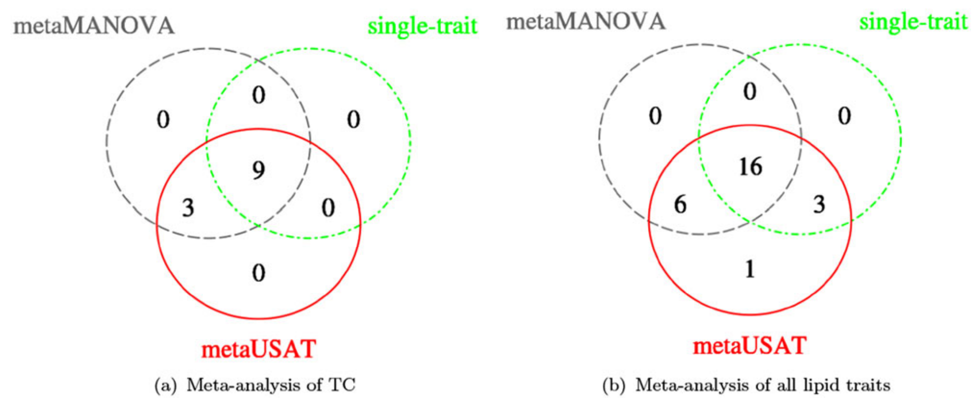


FIGURE 3 METSIM+T2D-GENES studies: Venn diagram of the number of SNPs (and not independent loci) found significant by each of metaUSAT, metaMANOVA, and single-trait analyses

Notes: A total of 31,897 SNPs (MAF ≥ 1%) are tested. For the single-trait analysis, a variant is declared as significant if its *P*-value for at least one trait is $< 1.25 \times 10^{-8}$ (four-trait Bonferroni corrected GWAS threshold). It should be noted that most of these significant SNPs are in LD.

TABLE 4 T2D-GENES + METSIM studies: Meta-analysis of all four lipid traits

			P-value		Empirical *P*-value	
rsID	chr	Position	meta-USAT	meta-MANOVA	metaUSAT	Known association result
rs2483205	1	55518316	$\mathbf{2.5 \times 10^{-8}}$	1.1×10^{-7}	$\mathbf{3.2 \times 10^{-8}}$	Lipids, Lipoprotein fractions[a]
rs1367117[b]	2	21263900	$\mathbf{1.5 \times 10^{-9}}$	2.3×10^{-7}	$\mathbf{2.8 \times 10^{-9}}$	Lipids, Lipoprotein fractions[c]
rs2304130	19	19789528	$\mathbf{1.5 \times 10^{-9}}$	1.8×10^{-6}	$\mathbf{3.3 \times 10^{-9}}$	Lipids, Lipoprotein fractions, T2D[d]

Notes: This table lists the SNPs exclusively detected by metaUSAT only. Only the independent SNPs (pairwise distance > 500 kb and $r^2 < 0.1$) are listed. We also report the empirical *P* values of metaUSAT based on 8.5×10^9 Monte Carlo simulations. *P* values exceeding the genome-wide threshold of 5×10^{-8} have been bold-faced. The known association results are based on previously reported GWAS associations within 500 kb of and $r^2 > 0.7$ with any of our SNPs from the NHGRI GWAS catalog (Welter et al., 2014) and our in-house GWAS catalog.

[a]Near many known GWAS hits for lipids (Surakka et al., 2015), lipoprotein fractions (Kettunen et al., 2012), and cardiovascular endpoints (Kathiresan et al., 2009).

[b]Illumina OmniExpress Exome Chip ID is exm176096.

[c]Known GWAS hit for lipids (Teslovich et al., 2010; Willer et al., 2013; Surakka et al., 2015)

[d]Known GWAS hit for lipids (Kristiansson et al., 2012; Willer et al., 2013)

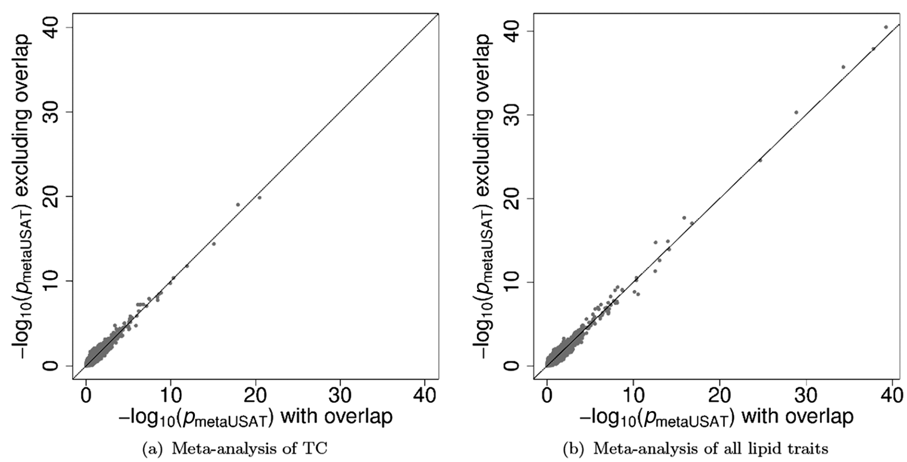


FIGURE 4 METSIM+T2D-GENES studies: metaUSAT *P*-values, with the overlapping individuals in the two studies, are plotted on the *x*-axis, while metaUSAT *P*-values after removing the overlap from METSIM are plotted on the *y*-axis

error for moderate α levels, but produce slightly inflated type I error rates at very small α levels (like the other methods). We found that metaMANOVA has a serious drawback: it may fail to detect association when most or all traits are associated (this behavior explored by Ray et al. (2016) in detail). The joint analysis of all lipid traits using METSIM and T2D-GENES studies further confirmed this. The power of metaMANOVA (and other multivariate tests) depends on a complex interplay of the number of truly associated traits, their correlation structure, and the directions of the signals. The underlying association scenario changes from one variant to another, and is not known a priori for any real dataset. There is no uniformly most powerful multivariate test, and a particular choice of association test may not be powerful enough to detect true signals. metaUSAT, being data-adaptive in nature, is less affected by the true (unknown) association scenario, and proves to be a robust yet computationally efficient choice for investigators.

The assumption of equal genetic effects across traits and studies is hardly tenable, making S_{Hom} unlikely to be powerful, especially when there is a moderate to large number of traits. aSPU relies on compute intensive *P*-value calculation approach, which is not feasible when analyzing large GWAS data. S_{Het} is usually dominated by metaMANOVA. On the other hand, metaUSAT is at least as powerful as metaMANOVA and a fast *P*-value calculation approach makes it suitable for testing genetic associations across multiple traits from multiple large-scale genome wide studies. Power of metaUSAT is robust to the proportion of associated traits. To alleviate any concern of inflated association signals of metaUSAT at stringent levels, we can calculate empirical metaUSAT *P* values (as described in supplementary S2). This need not be done for all variants; instead we can focus only on the handful of variants that have metaUSAT *P* values just crossing the chosen significance threshold.

metaUSAT can be used in a few different ways. We can test association of one or more traits from a single or multiple studies, which may or may not be independent. metaUSAT does not assume homogeneity of trait effects across studies. If the studies are nearly independent and the trait effects are believed to be homogeneous across studies, we can use meta-analyzed summary statistics for each trait (e.g., *Z*-statistic output from METAL; Willer, Li, & Abecasis, 2010) to perform joint meta-analysis of multiple traits. metaUSAT, also, does not require the independence of samples. When samples are related (e.g., in family-based GWAS), metaUSAT can use summary statistics from EMMAX (or other univariate mixed model framework) to appropriately test for genetic associations.

A potentially important contribution of metaUSAT can be in the emerging field of phenome-wide association studies (PheWAS) based on epidemiological cohorts. PheWAS systematically analyzes the impact of a genetic variant on a wide variety of human traits. Restrictions on data sharing necessitate use of meta-analysis for PheWAS (Bush, Oetjens, & Crawford, 2016). In this age of using publicly available data for increasing power and decreasing sequencing costs, overlapping samples may be a concern when it comes to meta-analysis. Furthermore, current single-trait meta-analysis approach for PheWAS is burdened by multiple comparison testing both at the variant level and at the trait level (Hebbring, 2014). We recommend using metaUSAT to overcome these challenges.

ACKNOWLEDGMENTS

The authors thank Anne Jackson for her help in obtaining summary statistics for the METSIM study. This research was supported by NIH grants HG000376 and DK062370 (MB).

WEB RESOURCES

We implemented metaUSAT in R. The software can be found in GitHub (https://github.com/RayDebashree/metaUSAT).

ORCID

Debashree Ray http://orcid.org/0000-0002-0979-2935

REFERENCES

Basu, S., Zhang, Y., Ray, D., Miller, M., Iacono, W., & McGue, M. (2013). A rapid gene-based genome-wide association test with multivariate traits. *Human Heredity*, *76*, 53–63.

Bolormaa, S., Pryce, J., Reverter, A., Zhang, Y., Barendse, W., Kemper, K., ... Goddard, M. (2014). A multi-trait, meta-analysis for detecting pleiotropic polymorphisms for stature, fatness and reproduction in beef cattle. *PLoS Genetics*, *10*, e1004198.

Bush, W. S., Oetjens, M. T., & Crawford, D. C. (2016). Unravelling the human genome-phenome relationship using phenome-wide association studies. *Nature Reviews Genetics*, *17*, 129–145.

Cichonska, A., Rousu, J., Marttinen, P., Kangas, A. J., Soininen, P., Lehtimäki, T., ... Pirinen, M. (2016). metaCCA: Summary statistics-based multivariate meta-analysis of genome-wide association studies using canonical correlation analysis. *Bioinformatics*, *32*, 1981–1989.

Conneely, K. N., & Boehnke, M. (2007). So many correlated tests, so little time! Rapid adjustment of p values for multiple correlated tests. *American Journal of Human Genetics*, *81*, 1158–1168.

Davies, R. (1980). Algorithm AS 155: The distribution of a linear combination of chi-square random variables. *Journal of the Royal Statistical Society. Series C, Applied Statistics*, *29*, 323–333.

Diggle, P., Heagerty, P., Liang, K.-Y., & Zeger, S. L. (2002). *Analysis of longitudinal data*. Oxford University Press, Oxford, UK.

Ferreira, M., & Purcell, S. (2009). A multivariate test of association. *Bioinformatics*, *25*, 132–133.

Fuchsberger, C., Flannick, J., Teslovich, T. M., Mahajan, A., Agarwala, V., Gaulton, K. J., ... McCarthy, M. I. (2016). The genetic architecture of type 2 diabetes. *Nature*, *536*, 41–47.

Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., & Hothorn, T. (2016). mvtnorm: Multivariate Normal and t Distributions. R package version 1.0-5.

He, L., Kernogitski, Y., Kulminskaya, I., Loika, Y., Arbeev, K., Loiko, E., ... Kulminski, A. M. (2016). Pleiotropic meta-analyses of longitudinal studies discover novel genetic variants associated with age-related diseases. *Frontiers in Genetics*, *7*, 179.

Hebbring, S. J. (2014). The challenges, advantages and future of phenome-wide association studies. *Immunology*, *141*, 157–165.

Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S.-y., Freimer, N. B., ... Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics*, *42*, 348–354.

Kathiresan, S., Voight, B. F., Purcell, S., Musunuru, K., Ardissino, D., Mannucci, P. M., ... Altshuler, D. (2009). Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nature Genetics*, *41*, 334–341.

Kettunen, J., Tukiainen, T., Sarin, A.-P., Ortega-Alonso, A., Tikkanen, E., Lyytikäinen, L.-P., ... Ripatti, S. (2012). Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nature Genetics*, *44*, 269–276.

Kim, J., Bai, Y., & Pan, W. (2015). An adaptive association test for multiple phenotypes with GWAS summary statistics. *Genetic Epidemiology*, *39*, 651–663.

Kristiansson, K., Perola, M., Tikkanen, E., Kettunen, J., Surakka, I., Havulinna, A. S., ... Salomaa, V. (2012). Genome-wide screen for metabolic syndrome susceptibility loci reveals strong lipid gene contribution but no evidence for common genetic basis for clustering of metabolic syndrome traits. *Circulation: Cardiovascular Genetics*, *5*, 242–249.

Lin, D.-Y., & Sullivan, P. (2009). Meta-analysis of genome-wide association studies with overlapping subjects. *American Journal of Human Genetics*, *85*, 862–872.

Liu, H., Tang, Y., & Zhang, H. (2009). A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables. *Computational Statistics & Data Analysis*, *53*, 853–856.

Liu, Z., & Lin, X. (2017). Multiple phenotype association tests using summary statistics in genome-wide association studies. *Biometrics*. https://doi.org/10.1111/biom.12735.

Majumdar, A., Witte, J. S., & Ghosh, S. (2015). Semiparametric allelic tests for mapping multiple phenotypes: Binomial regression and Mahalanobis distance. *Genetic Epidemiology*, *39*, 635–650.

O'Brien, P. (1984). Procedures for comparing samples with multiple endpoints. *Biometrics*, *40*, 1079–1087.

Pan, W. (2009). Asymptotic tests of association with multiple SNPs in linkage disequilibrium. *Genetic Epidemiology*, *33*, 497–507.

Pausch, H., Emmerling, R., Schwarzenbacher, H., & Fries, R. (2016). A multi-trait meta-analysis with imputed sequence variants reveals twelve QTL for mammary gland morphology in Fleckvieh cattle. *Genetics Selection Evolution*, *48*. https://doi.org/10.1186/s12711-016-0190-4

Porter, H., & O'Reilly, P. (2017). Multivariate simulation framework reveals performance of multi-trait GWAS methods. *Scientific Reports*, *7*. https://doi.org/10.1038/srep38837.

Ray, D., & Basu, S. (2017). A novel association test for multiple secondary phenotypes from a case-control GWAS. *Genetic Epidemiology*, *41*, 413–426.

Ray, D., Li, X., Pan, W., Pankow, J. S., & Basu, S. (2015). A Bayesian partitioning model for detection of multilocus effects in case-control studies. *Human Heredity*, *79*, 69–79.

Ray, D., Pankow, J. S., & Basu, S. (2016). USAT: A unified score-based association test for multiple phenotype-genotype analysis. *Genetic Epidemiology*, *40*, 20–34.

Stancčáková, A., Javorský, M., Kuulasmaa, T., Haffner, S., Kuusisto, J., & Laakso, M. (2009). Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6416 Finnish men. *Diabetes*, *58*, 1212–1221.

Stephens, M. (2013). A unified framework for association analysis with multiple related phenotypes. *PLoS One*, *8*, e65245.

Surakka, I., Horikoshi, M., Mägi, R., Sarin, A.-P., Mahajan, A., Lagou, V., ... ENGAGE Consortium. (2015). The impact of low-frequency and rare variants on lipid levels. *Nature Genetics*, *47*, 589–597.

Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., ... Kathiresan, S. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*, *466*, 707–713.

Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., ... Parkinson, H. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Research*, *42*, D1001–D1006.

Willer, C. J., Li, Y., & Abecasis, G. R. (2010). METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, *26*, 2190–2191.

Willer, C. J., Schmidt, E. M., Sengupta, S., Peloso, G. M., Gustafsson, S., Kanoni, S., ... Global Lipids Genetics Consortium. (2013). Discovery and refinement of loci associated with lipid levels. *Nature Genetics*, *45*, 1274–1283.

Wu, B., & Pankow, J. S. (2016). On sample size and power calculation for variant set-based association tests. *Annals of Human Genetics*, *80*, 136–143.

Wu, M., Lee, S., Cai, T., Li, Y., Boehnke, M., & Lin, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. *American Journal of Human Genetics*, *89*, 82–93.

Zhou, X., & Stephens, M. (2014). Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods*, *11*, 407–409.

Zhu, X., Feng, T., Tayo, B. O., Liang, J., Young, J. H., Franceschini, N., ... Redline, S. (2015). Meta-analysis of correlated traits via summary statistics from GWASs with an application in hypertension. *American Journal of Human Genetics*, *96*, 21–36.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Ray D, Boehnke M. Methods for meta-analysis of multiple traits using GWAS summary statistics. *Genet Epidemiol*. 2018;42:134–145. https://doi.org/10.1002/gepi.22105