Published in final edited form as:

Genet Epidemiol. 2018 December ; 42(8): 812–825. doi:10.1002/gepi.22148.

Adjustment for Covariates Using Summary Statistics in GWAS analysis

Tao Wang*, Xiaonan Xue*, Xianhong Xie*, Kenny Ye*, Xiaofeng Zhu⁺, and Robert C. Elston⁺

*Department of Epidemiology & Population Health, Albert Einstein College of Medicine, Bronx, New York, United States of America

⁺Department of Population and Quantitative Health Sciences, School of Medicine, Case Western Reserve University, Cleveland, OH, United States of America

Abstract

Linear regression is a standard approach to identify genetic variants associated with continuous traits in genome-wide association studies (GWAS). In a standard epidemiology study, linear regression is often performed with adjustment for covariates to estimate the independent effect of a predictor variable or to improve statistical power by reducing residual variability. However, it is problematic to adjust for heritable covariates in genetic association analysis. Here, we propose a new method that utilizes summary statistics of the covariate from additional samples for reducing the residual variability and hence improves statistical power. Our simulation study showed that the proposed methodology can maintain a good control of type I error and can achieve much higher power than a simple linear regression. The method is illustrated by an application to the GWAS results from the GIANT consortium.

Keywords

GWAS; Meta-analysis; Linear Regression; Pleiotropy; Continuous traits

Introduction

Genome-wide association studies (GWAS), whole-exome sequencing studies (WES) and whole genome sequencing studies (WGS) have produced many large datasets of genetic variation in hundreds of thousands to millions of human subjects (Boyle, Li, & Pritchard, 2017), which has motivated the development of new statistical methods to analyze these

To whom correspondence should be addressed at: tao.wang@einstein.yu.edu.

Supplemental Data

Supplemental data include one table.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Web Resources

The URLs for data presented herein are as follows:

GIANT consortium data files, <http://www.broadinstitute.org/collaboration/giant>

OMIM, <http://www.omim.org/>

particular because it may miss important genetic variants with pleiotropic effects. Furthermore, this method requires individual level data for all the covariates.

In recent years, GWAS summary statistics of various heritable covariates have become available in many large studies. For example, GWAS of hundreds of thousands subjects have already been performed for BMI, height, and blood pressure that are commonly-used as covariates for many complex diseases such as cancers, heart vascular diseases, and diabetes (Locke et al., 2015; Warren et al., 2017; Wood et al., 2014). In this report, we developed Adjustment Analysis using Meta-statistic of Covariates (*AAMC*), a new covariate adjustment method to improve statistical power by taking advantage of these existing GWAS results of covariate when identifying genetic variants associated with continuous traits. Instead of only including covariates that are not associated with genetic variants, the proposed approach is to effectively remove any potential correlation between the trait and covariates due to the genetic variant of interest. To do this, a naive approach is to adjust for the residual obtained by regressing the covariate on the genetic variant. Because the residuals are independent of the genetic variant, the estimate of genetic effect on the trait is unbiased. In general, regressing an independent variable on the residual of the outcome adjusted for a covariate can improve statistical power by reducing the error variance. However, adding the residual of the covariate adjusted for a SNP can induce a “residual” bias because it fails to account for the error associated with estimating the effect of the SNP on this covariate (Aschard, Vilhjalmsen, Joshi, Price, & Kraft, 2015). After appropriately accounting for the estimation variability, adjustment for the residual of the covariate does not improve statistical power (see Method section). The idea of *AAMC* is to make use of additional information from existing GWAS summary statistics of the covariate to reduce the variance of the estimated genetic effect on the covariate; thereby, it can substantially improve statistical power while maintaining type I error. The proposed method is first derived assuming individual data on the trait of interest are available. We then extend the method based on having only GWAS summary statistics. We evaluated the *AAMC* approaches by analyzing simulation data and summary statistics of WC and WHR from the GIANT consortium.

Method

Notation

Consider a GWAS of n subjects. For subject i ($i=1, \dots, n$), let x_i be the genotype value of a SNP, y_i be a continuous trait, and c_i be a continuous covariate correlated with the trait. Without loss of generality, we assume y_i and c_i are standardized. We consider a pleiotropic model in which y_i and c_i are associated with the SNP by:

$$\begin{aligned} y_i &= u_i + ax_i + w_{yi} \\ c_i &= u_i + bx_i + w_{ci} \end{aligned} \quad (1)$$

where u_i is a random variable with zero mean and variance σ_u^2 , w_{yi} and w_{ci} are independent residuals with $\text{Var}\begin{pmatrix} w_y \\ w_c \end{pmatrix} = \begin{bmatrix} \sigma_y^2 & 0 \\ 0 & \sigma_c^2 \end{bmatrix}$, and a and b are genetic effects on the trait and covariate, respectively. Under this model, the correlation of y_i and c_i due to unidentified common environment and genetic effects is modeled by the random effect u_i and

$$\text{Var}\begin{pmatrix} y_i \\ c_i \end{pmatrix} = \begin{bmatrix} \sigma_u^2 + b^2\sigma_x^2 + \sigma_y^2 & a\sigma_x^2 \\ a\sigma_x^2 & \sigma_x^2 \\ \sigma_u^2 + b^2\sigma_x^2 + \sigma_c^2 & b\sigma_x^2 \\ b\sigma_x^2 & \sigma_x^2 \end{bmatrix}.$$

Bias of adjustment for the covariate

Our goal is to detect the association between x and y . A regression model including the covariate c is given by

$$E(y_i | x_i, c_i) = \gamma c_i + \tilde{a} x_i, \quad (2)$$

where γ and $\tilde{a}$ are regression coefficients for the covariate and the SNP. The effect estimate of the SNP on the trait is given by

$$\tilde{a} = \frac{\sum_i (c_i^2) \sum_i (x_i y_i) - \sum_i (x_i c_i) \sum_i (y_i c_i)}{\sum_i (c_i^2) \sum_i (x_i^2) - (\sum_i (x_i c_i))^2}.$$

Dividing numerator and denominator by n^2 and letting n tend to infinity,

$$\tilde{a} = \frac{E(c^2)E(xy) - E(xc)E(y_c)}{E(c^2)E(x^2) - [E(xc)]^2}.$$

Under the null hypothesis of $a = 0$,

$$\tilde{a} \rightarrow \frac{-\text{Cov}(x, c)\text{Cov}(y, c)}{\text{Var}(c)\text{Var}(x) - [\text{Cov}(x, c)]^2} = \frac{-b\sigma_x^2\sigma_u^2}{(\sigma_u^2 + b^2\sigma_x^2 + \sigma_c^2)\sigma_x^2 - (b\sigma_x^2)^2} = \frac{-b\sigma_u^2}{\sigma_u^2 + \sigma_c^2}.$$

Because the genetic effect of a SNP is often small, $\tilde{a} \approx -\frac{br_{y,c}\sigma_y\sigma_c}{\sigma_c^2} = -\frac{br_{y,c}\sigma_y}{\sigma_c} = -br_{y,c}$, where $r_{y,c}$ is the correlation

between y and c . Hence, when the covariate is associated with the SNP ($b \neq 0$) and the trait and covariate are correlated ($\sigma_u^2 \neq 0$), $E(\tilde{a}) \neq 0$, leading to a spurious association.

Adjusting for residuals of the covariate

The above equation suggests an unbiased result may be obtained by adjusting out the covariate *after* removing the SNP effect on the covariate. To do this, a naïve idea is to adjust out the covariate residuals. Let the covariate residual $\tau_i = c_i - \hat{\beta}x_i$ in which $\hat{\beta}$ is estimated from the regression model $E(c_i) = \hat{\beta}x_i$. On replacing the covariate by its residual in regression (2), we have a regression model of the form

$$E(y_i | x_i, \tau_i) = \gamma\tau_i + \alpha x_i. \quad (3)$$

Under this model, we have

$$\hat{\alpha} = \frac{\sum_i (x_i y_i)}{\sum_i (x_i^2)}.$$

This estimate is the same as that of the simple regression model of $E(y_i) = \alpha x_i$ and is unbiased. Given $\hat{b}$,

$$\text{Var}(\hat{\alpha} | \hat{b}) = \sigma_{y, \tau}^2 / \sum_i (x_i^2),$$

where $\sigma_{y, \tau}^2$ is the residual variance of regression model adjusting for τ (3). The variance of the simple regression is $\text{Var}(\hat{\alpha}) = \sigma_y^2 / \sum_i (x_i^2)$. By comparing $\text{Var}(\hat{\alpha})$ and $\text{Var}(\hat{\alpha} | \hat{b})$, we have $\frac{\text{Var}(\hat{\alpha})}{\text{Var}(\hat{\alpha} | \hat{b})} = 1/(1 - r_{y, c}^2)$; that is, the variance is underestimated due to the fact that $\hat{b}$ is estimated from the same data as used for testing the association between the trait and the SNP. By accounting for the variability of $\hat{b}$, the real variance of $\hat{\alpha}$ is

$$\text{Var}(\hat{\alpha}) = E[\text{Var}(\hat{\alpha} | \hat{b})] + \text{Var}[E(\hat{\alpha} | \hat{b})],$$

where $E[\text{Var}(\hat{\alpha} | \hat{b})] = \text{Var}(\hat{\alpha} | \hat{b})$, $\text{Var}[E(\hat{\alpha} | \hat{b})] = r_{\hat{\alpha}, \hat{b}}^2 \text{Var}(\hat{\alpha})$, $r_{\hat{\alpha}, \hat{b}}$ is the correlation between $\hat{\alpha}$ and $\hat{b}$. Hence, $\text{Var}(\hat{\alpha}) = \text{Var}(\hat{\alpha} | \hat{b}) / (1 - r_{\hat{\alpha}, \hat{b}}^2)$. Let $X = (x_1, x_2, \dots, x_n)^T$ and $G = (X^T X)^{-1} X^T$.

When the same set of samples is used for estimating both a and b , $\hat{a} = Gy$ and $\hat{b} = Gc$,

$$r_{\hat{a}, \hat{b}} = \frac{G^T \text{Cov}(y, c) G}{\sqrt{G^T G \sigma_y^2} \sqrt{G^T G \sigma_c^2}} = r_{y, c}.$$

Thus $Var(\hat{a}) = \frac{Var(\hat{a} | \hat{b})}{1 - r_{\hat{a}, \hat{b}}^2} = \frac{Var(\hat{a})(1 - r_{y, c}^2)}{1 - r_{\hat{a}, \hat{b}}^2} = Var(\hat{a})$, suggesting that adjustment for residuals

does not improve the power of simple regression.

Incorporating GWAS summary statistics of the covariate

However, suppose the samples used for estimating $\hat{b}$ and $\hat{a}$ are not completely overlapped, let N_a be the sample size for estimating $\hat{a}$, N_b the sample size for estimating $\hat{b}$, and N_s the sample size used for estimating both $\hat{a}$ and $\hat{b}$. It can be shown that

$$r_{\hat{a}, \hat{b}} = \frac{N_s}{\sqrt{N_a N_b}} r_{y, c}.$$

This equation indicates $|r_{\hat{a}, \hat{b}}| < |r_{y, c}|$ when there are additional samples for estimating of b ($N_b > N_a$). In this case, the variance of the SNP effect after adjustment for the residuals is less than that without adjustment for the residuals, i.e. $Var(\hat{a}) < Var(\hat{a})$.

This suggests that the statistical power of testing association between the trait and the SNP can be improved by adjusting for the residual if we incorporate the GWAS result of the covariate from additional samples. Let $\hat{b}_0$ ($SE(\hat{b}_0)$) and $\hat{b}_1$ ($SE(\hat{b}_1)$) be the effect size estimate (standard error) from additional samples with a sample size n_0 and the current samples with a sample size n_1 , respectively. Assuming the two sets of samples are from the same population, the information can be combined by a fixed-effect meta-analysis in which

$$\hat{b}_{meta} = \frac{\sum_k (w_k b_k)}{\sum_k w_k}$$

$$SE(\hat{b}_{meta}) = \frac{1}{\sqrt{\sum_k w_k}},$$

where $w_k = 1/SE(\hat{b}_k)^2$ for $k = 0, 1$ for the additional and current samples, respectively. Let $\tau_{i, meta} = c_i - \hat{b}_{meta} x_i$. Replacing τ_i by $\tau_{i, meta}$ in regression (3), the new regression model is

$$E(y_i | x_i, \tau_{i, meta}) = \check{\gamma}_{meta} \tau_{i, meta} + \check{a}_{meta} x_i. \quad (4)$$

Let $\hat{a}_{meta}$ and $Var(\hat{a}_{meta} | \hat{b}_{meta})$ be the estimates of the genetic effect and corresponding variance from regression model (4). As discussed above, $Var(\hat{a}_{meta} | \hat{b}_{meta})$ underestimates the actual variance because the variability of $\hat{b}_{meta}$ is not taken into account in the variance estimate. Under the null hypothesis $a = 0$, it can be shown that the actual variance is given by

$$Var(\hat{a}_{meta}) \approx Var(\hat{a}_{meta} | \hat{b}_{meta}) + Var(\hat{b}_{meta})\hat{r}_{y,c}^2$$

(Appendix 1), in which $\hat{r}_{y,c}$ can be estimated from the SNP summary statistics of the trait and the covariate obtained from a set of independent SNPs across the genome after excluding the variants associated with the trait and the covariate (Park, Li, Song, He, & Zhu, 2016), or from the LD score regressions (Bulik-Sullivan et al., 2015). We can define the test statistic for the SNP-association of the trait as

$$Q = \frac{\hat{a}_{meta}^2}{Var(\hat{a}_{meta})},$$

and Q can be approximated by a $\chi_{df=1}^2$ distribution under the null hypothesis.

In the above, we require individual data for calculating residuals to perform regression model (4). However, when only summary statistics with adjustment for the covariate [$\hat{a}$ and $Var(\hat{a})$] are available, we can directly correct for the bias of $\hat{a}$ using $\hat{b}_{meta}$. Because the bias is $-br_{y,c}$, the bias corrected genetic effect on the trait can be estimated by $\hat{a} + \hat{b}_{meta}\hat{r}_{y,c}$. It can be shown that $Var(\hat{a} + \hat{b}_{meta}\hat{r}_{y,c}) = Var(\hat{a}) + \hat{r}_{y,c}^2 Var(\hat{b}_{meta})$ (Appendix 2). We define the test statistic based on summary statistics as

$$Q_{hom} = \frac{(\hat{a} + \hat{b}_{meta}\hat{r}_{y,c})^2}{Var(\hat{a}) + \hat{r}_{y,c}^2 Var(\hat{b}_{meta})},$$

The distribution of Q_{hom} can be approximated by a $\chi_{df=1}^2$ distribution under the null hypothesis.

Heterogeneity of SNP effects on the covariate

Above, we assumed that a previous GWAS for estimating the genetic effect on the covariate was done in the same population of the current study. To allow for potential heterogeneity in genetic effects on the covariate between the previous and current studies, we can consider a random effects model:

$$\hat{b}_k = \mu_b + \delta_k + e_k,$$

where μ_b is the mean effect, δ_k is the random effect associated with study k , and e_k is a random error. Let $S_k = Var(e_k)$, which is available from each study. If the sample size is sufficiently large, e_k is approximately $N(0, S_k)$. We assume that $\delta_k \sim N(0, D)$. Under this model, the heterogeneity between the two studies is modelled by the random effect δ_k . The

parameters of the model (D and μ_b) can be estimated by a generalized linear squares approach for meta-analysis (Berkey, Hoaglin, Antczak-Bouckoms, Mosteller, & Colditz, 1998). By incorporating information on b from the samples in the previous study, we can estimate the best linear unbiased predictor (BLUP) of the genetic effect on the covariate in the current study (b_1) by

$$\hat{b}_{BLUP} = \lambda \hat{b}_1 + (1 - \lambda) \hat{\mu}_b,$$

where $\lambda = \hat{D}/(\hat{D} + S_1)$. This estimate is intuitively reasonable because, when there is no or little heterogeneity, $\hat{b}_{BLUP}$ is close to μ_b (similar to the estimate from a fixed effect model) and it borrows information from the additional sample in the previous study on b for improving statistical power. However, when there is significant heterogeneity between the two studies, it tends to be close to $\hat{b}_1$ for avoiding bias. By replacing $\hat{b}_{meta}$ by $\hat{b}_{BLUP}$, we have a BLUP version of AAMC:

$$Q_{het} = \frac{(\hat{a} + \hat{b}_{BLUP} \hat{r}_{y,c})^2}{Var(\hat{a}) + \hat{r}_{y,c}^2 Var(\hat{b}_{BLUP})}.$$

This BLUP version of the statistic can also be approximated by a $\chi^2_{df=1}$ distribution under the null hypothesis.

Results

Simulation Design

We examined the proposed method by simulations to assess the type I error and power. In each replicate of our simulations, we first generated the genotype values of a SNP by sampling from a Binomial distribution with a specified allele frequency. The minor allele frequency of the SNP was set at 0.1. Then, the trait and the covariate were simulated based on model (1), where the random errors w_{yi} and w_{ci} were sampled from a standard normal distribution, and the random effect u_i was sampled from a normal distribution with mean zero and various standard deviations to generate different correlations between the trait and the covariate. Different SNP effect sizes (b) on the covariate were considered for generating various levels of bias in regression analysis with adjustment for the covariate. Two groups of subjects were generated for the trait analysis and the additional covariate analysis, respectively. Both the trait and covariate were assumed to be available for the trait analysis, and the sample size was set at 1,000 (n_1). Only the covariate measurement was assumed to be available for the additional covariate analysis, and various additional sample sizes ($n_0=1000, n_0=2000, 5000$, and $10,000$) were considered. For both Q_{hom} and Q_{het} , summary statistics were first calculated based on the individual data of the two groups.

We first performed a simulation study to confirm that the proposed methods can maintain a good control of the type I error. For evaluating the type I error, the genetic effect a was set to zero. We first considered a homogeneous model, in which the SNP effect on the covariate in

the two groups was the same ($b_0 = b_1$). We then considered a heterogeneous model. We used a method similar to that used before for simulating heterogeneous effects (Nakaoka & Inoue, 2009). Specifically, the SNP effect on the covariate was sampled from $N(b, \sigma_b^2)$, where σ_b^2 is the between-study variance and the values of σ_b varied from small (0.01) to large (0.1) (Moonesinghe, Khoury, Liu, & Ioannidis, 2008). For each scenario, we simulated 10,000 repeats and estimated the empirical type I error by the proportion of repetitions with p-values greater than 0.01.

For evaluating statistical power, we considered various proportions of trait variation explained by the SNP by assuming different SNP effects on the trait (a). We considered two scenarios: (1) a model without pleiotropic effects ($a > 0$ and $b = 0$) and (2) a model with pleiotropic effects ($a > 0$ and $b > 0$). We evaluated statistical power at the significance level of 0.01 using 1,000 replicates.

Type I error

Table 1 shows the empirical type I error of the different tests at the nominal type I error level 0.01 under a homogeneous model in which the SNP had the same effect on the covariate in samples for the trait analysis and the additional covariate analysis ($b_0 = b_1 = b$). We can see that the empirical type I error of the proposed three tests (Q , Q_{hom} and Q_{het}) was around the nominal level, indicating that our approximation of the null distribution is adequate at the 0.01 significance level. As expected, the regression model with adjustment for the covariate (Adj) could not maintain control of the type I error, which was inflated when there is a correlation between the trait and the covariate (i.e. $\sigma_u^2 > 0$) and the SNP effect on the covariate is not zero (i.e. $b \neq 0$). Also, the type I error of the regression model with adjustment for the covariate residual (Adj_τ) increased with increasing correlation between the trait and the covariate (σ_u^2) because it did not account for variability of the regression coefficient estimate of the SNP effect on the covariate.

Table 2 shows the empirical type I error of the proposed tests at the nominal Type I error level 0.01 under a heterogeneous model, in which ($b_0 \neq b_1$). In these simulations, b_0 and b_1 were independently sampled from $N(0.2, \sigma_b^2)$. When the sample size of the additional study of the covariate was similar to that of the trait analysis ($n_0 = n_1 = 1,000$), Q and Q_{hom} still had a robust type I error. However, the type I error of Q and Q_{hom} tended to increase with an increasing sample size of the group for the covariate analysis (e.g. $n_1 = 1,000$ and $n_0 = 5,000$), especially when the heterogeneity level was high ($\sigma_b = 0.1$). This result is expected because, when n_0 is much larger than n_1 , the estimate of the SNP effect on the covariate from the meta-analysis based on a fixed effect model is dominated by the estimate from the group for the covariate analysis, and does not reflect the true SNP effect on the covariate in the group for the trait analysis. Nevertheless, the empirical type I error of Q_{het} was still around the nominal level 0.01, even when the additional sample size was large ($n_0 = 5,000$) and the heterogeneity was high ($\sigma_b = 0.1$), because the BLUP estimate accounts for the heterogeneity in SNP effects as well as the difference in sample sizes.

Power

Figure 1 shows the empirical power of different tests for testing the association of the trait with a SNP at the significance level 0.01 under various sample sizes of the additional group for the covariate analysis under a model without pleiotropic effects ($b = 0$). The proposed tests (Q , Q_{hom} and Q_{het}) clearly outperformed the simple regression, because they were able to reduce residual variance of the regression by making use of information on the covariate from the additional study. Given the same correlation between the trait and covariate, power increased when the sample size of the additional study for the covariate increased, because of an improved estimate of the SNP effect on the covariate. As expected, the regression model with adjustment for the covariate had the best performance in the setting that the SNP does not have any effect on the covariate. In this situation, the proposed tests were less powerful than the regression with adjustment for the covariate, because they involved additional variability from estimating the SNP effect on the covariate. However, the loss in power of the proposed tests tended to be small with increasing sample size of the additional study on the covariate. The performance of Q_{hom} was slightly better than Q_{het} because the variance of the SNP effect estimate on the covariate is in general larger than that of Q_{hom} for accounting for heterogeneity between studies.

Figure 2 shows the empirical power of different tests for testing the association of the trait with a SNP at the 0.01 significance level with various correlations between the trait and covariate under a model without pleiotropic effects. In our simulations, the power of a simple regression significantly decreased with an increasing covariance between the trait and covariate (σ_u^2), because the variance of the trait explained by the SNP decreased with fixed SNP effects. However, the proposed tests were relatively less impacted by increasing σ_u^2 , and hence the gain in power of the proposed test is more obvious when the correlation between the trait and covariate is large. As expected, the performance of the regression with adjustment for the covariate was robust to increasing σ_u^2 because the adjustment can completely remove the correlation between the trait and covariate.

Figures 3 and 4 show the empirical power of different tests for testing the association of the trait with a SNP at the 0.01 significance level when the SNP has pleiotropic effects on both the trait and covariate. In most situations, the proposed tests had the best performance. The regression with adjustment for the covariate, as expected, performed very poorly unless the SNP effect was very large, because the negative bias due to adjustment for the covariate masked the SNP effect on the trait.

Analyzing WHR and waist in the GIANT consortium

To illustrate the utility of the approach proposed here, we have applied the method to WC and WHR summary statistics from GIANT. The data from the 2015 meta-analysis were used because all summary statistics for BMI, WC and WHR, as well as for WC adjusted for BMI (WCadj) and WHR adjusted for BMI (WHRadj) were available (<http://portals.broadinstitute.org/collaboration/giant>). For reducing the heterogeneity in the SNP effect, only European ancestry data were included in the analysis. We assumed a homogeneous model in which the SNP effects on BMI are similar between subjects with

both BMI and WC (or WHR) measurements and subjects with only BMI measurements. The median sample sizes were about 230,000, 140,000 and 150,000 for BMI, WHR and WC, respectively. With the larger sample size for BMI, we could improve the power of the regression analysis to identify SNPs associated with WC (or WHR) by adjusting for BMI while maintaining control of type I error in the presence of pleiotropic effects on BMI and WC (or WHR). We used previously calculated correlations in the GIANT dataset, 0.85 and 0.49, between BMI and WC and between BMI and WHR, respectively). (Aschard et al., 2015) (Heid et al., 2010).

As shown by Figure 5 (a), the proposed method (Q_{hom}), the regression with adjustment for BMI (Adj), and the regression without adjustment (Noadj) identified 780, 717 and 812 SNPs, respectively, associated with WC with P -value $< 10^{-8}$. Most of these SNPs are in regions of high linkage disequilibrium (LD). Among the 780 SNPs identified by Q_{hom} , 574 SNPs were not identified by Adj (P -value $> 10^{-8}$), and 467 SNPs did not even achieve the nominal significance level 0.01. The results of Noadj indicated all 467 SNPs were associated with both BMI and WC (P -value $< 10^{-8}$). This is consistent with our simulation results that Adj is likely to miss SNPs with pleiotropic effects on both the trait and the covariate. Among the 717 SNPs identified by Adj, 511 SNPs were not significant for the proposed Q_{hom} (P -value $> 10^{-8}$), but only 23 SNPs in 7 regions of high LD did not achieve the nominal significance level 0.01. For the 23 SNPs, the effect estimates on WC using Adj were in the opposite direction to their effect estimates on BMI (Supplement table 1), suggesting they may be spurious associations because of the negative bias from the adjustment for the heritable trait BMI. The Noadj method identified the most SNPs with P -value $< 10^{-8}$ because it likely also captures SNPs with an indirect effect through BMI. Figure 5 (b) plots the Q_{hom} P -values against regression with and without adjustment for BMI. High discordance of P -values between Q_{hom} and Adj, in particular for SNPs significantly associated with WC by Adj, suggests that there may be a large proportion of false significant SNPs identified by Adj.

As shown by Figure 6 (a), the proposed Q_{hom} and regressions with and without adjustment for BMI identified 415, 618 and 399 SNPs, respectively, associated with WHR with P -value $< 10^{-8}$. Among the 415 SNPs identified by Q_{hom} , regression with adjustment did not identify 67 SNPs (P -value > 0.01), which may have pleiotropic effects that led to a negative bias and hence reduced power of the regression with adjustment. These 67 SNPs were associated with both BMI and WHR (P -value $< 10^{-8}$) by simple regression without adjustment, which also suggests potential pleiotropic effects. In contrast to WC, among the 618 SNPs identified by regression with adjustment, 270 SNPs were not identified by Q_{hom} ($P > 10^{-8}$), but all 270 SNPs achieved the nominal significance level of 0.01 by both Q_{hom} and Noadj, suggesting that the bias of Adj in the WHR analysis might be less severe than that in WC. This may be due to the fact that the correlation between WHR and BMI was less strong than that between WC and BMI (0.49 vs 0.85). Figure 6 (b) plots the Q_{hom} P -values against regression with and without adjustment for BMI. Compared with Figure 5 (b), P -values between Q_{hom} and Adj are less discordant, also suggesting that the bias of Adj might be less severe in the WHR analysis.

Discussion

In the present study, we developed a new method under the framework of a linear regression model to utilize additional genetic information on a covariate for improving the power of regression analysis, while avoiding an inflated type I error. It has been shown that artificial association can be created by the bias of linear regression when a heritable covariate is included as a covariate (Aschard et al., 2015; Day, Loh, Scott, Ong, & Perry, 2016). The bias can also lead to significant loss in statistical power when the SNP has pleiotropic effects on both the trait and covariate. This bias can be corrected by including the residual of the covariate instead of the covariate itself in the regression model, because the residual is not associated with the tested SNP. However, adjustment using residuals does not gain any power. Essentially, the regression with adjustment for covariate assumes that there is no SNP effect on the covariate, while regression with adjustment for residuals requires estimation of the SNP effect from the data, hence involving the additional variability of estimation. Thus there is a trade-off between bias (adjustment for covariate) and the loss in power (adjustment for residuals). However, there are existing GWAS results for many important covariates from many large cohorts. One may have to borrow available information on the covariate to improve statistical power while avoiding the bias. The proposed AAMC methods in this report naturally incorporate such prior information using a standard meta-analysis approach for reducing estimation variability of the SNP effect on the covariate. In simulations, we showed that the proposed methods can maintain good control of type I error and substantially improve power over conventional methods.

In general, the effects of genetic variants on complex diseases such as cancers tend to be relatively small. Thus current GWAS rely on very large sample sizes to achieve adequate statistical power, which has practical limitations. The proposed method can leverage the large amount existing genetic data on heritable covariates to gain statistical power. With the growing collection of data from comparable GWAS for commonly-used covariates such as BMI, height, and blood pressure, the proposed method can potentially have a wide use to identify genetic associations with complex diseases. Another advantage of the proposed AAMC is that it only requires summary statistics from GWAS and are simple to compute; the method uses existing meta-analysis approaches to estimate the effect of a SNP on the covariate, which is available in many software packages. Further, the test statistic can be asymptotically approximated by a standard chi-square distribution, and hence the p-value can be evaluated quickly and simply, avoiding extensive simulation to evaluate the empirical null distribution.

The proposed method has its limitations. Our method uses information on the covariate from other GWAS to improve the power. The method relies on the accuracy of the estimate of the SNP effect on the covariate, obtained through the meta-analysis. Consequently, the result of the proposed method can be sensitive to heterogeneity between the main study of the trait analysis and the additional studies of the covariate analysis. In practice, it is important to identify comparable studies with similar genetic and environmental backgrounds, genotyping technology, and measurement techniques for the covariate. In addition, it may also be helpful to check if the SNP effect on the covariate is homogeneous between the main study and the additional study before the analysis. There are several sources of heterogeneity

of the SNP effect. Major sources of heterogeneity may be gene-gene and gene-environment interactions. When two groups of individuals are sampled in populations with varying genetic and environmental backgrounds, the effect size of a genetic variant varies depending on the genetic and environmental backgrounds (Hirschhorn & Daly, 2005; Hunter, 2005; Ioannidis, Patsopoulos, & Evangelou, 2007; Moonesinghe et al., 2008). Heterogeneity can also occur because traits are measured differently. A third source of heterogeneity is due to the use of different GWAS panels, resulting in the genotypes of some SNPs being imputed. As found in our simulations, the type I error of the proposed methods (Q and Q_{hom}) depends on the level of heterogeneity. It may be difficult to completely remove all sources of heterogeneity. Hence, it is important to have a statistical method that is robust to heterogeneity at a reasonable level. In this paper, we suggest using the BLUP estimate of the SNP effect to account for heterogeneity across studies. Our simulations indicated the proposed Q_{het} was more robust to heterogeneity than Q and Q_{hom} . It maintained the type I error under a wide range of heterogeneity levels, but is only slightly less powerful than the others.

In summary, we propose a method to improve the power of identifying a SNP association with a continuous trait by incorporating information on covariates from additional studies. The new test is simple to use, and hence feasible for large scale genetic studies with millions of genetic markers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported, in part, by R21 CA202529 (to T.W.) and by HG003054 (to X.F.Z.).

Grant numbers:

This research was supported, in part, by R21 CA202529 (to T.W.) and by HG003054 (to X.F.Z.).

References

- Aschard H, Guillemot V, Vilhjalmsen B, Patel CJ, Skurnik D, Ye CJ, ... Zaitlen N. Covariate selection for association screening in multiphenotype genetic studies. *Nat Genet.* 2017; 49(12): 1789–1795. DOI: 10.1038/ng.3975 [PubMed: 29038595]
- Aschard H, Vilhjalmsen BJ, Joshi AD, Price AL, Kraft P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am J Hum Genet.* 2015; 96(2):329–339. DOI: 10.1016/j.ajhg.2014.12.021 [PubMed: 25640676]
- Berkey CS, Hoaglin DC, Antczak-Bouckoms A, Mosteller F, Colditz GA. Meta-analysis of multiple outcomes by regression with random effects. *Stat Med.* 1998; 17(22):2537–2550. [PubMed: 9839346]
- Bhattacharjee S, Rajaraman P, Jacobs KB, Wheeler WA, Melin BS, Hartge P, ... Chatterjee N. A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. *Am J Hum Genet.* 2012; 90(5):821–835. DOI: 10.1016/j.ajhg.2012.03.015 [PubMed: 22560090]
- Boyle EA, Li YI, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell.* 2017; 169(7):1177–1186. DOI: 10.1016/j.cell.2017.05.038 [PubMed: 28622505]

- Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, ... Neale BM. Schizophrenia Working Group of the Psychiatric Genomics C. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet.* 2015; 47(3):291–295. DOI: 10.1038/ng.3211 [PubMed: 25642630]
- Chung D, Yang C, Li C, Gelernter J, Zhao H. GPA: a statistical approach to prioritizing GWAS results by integrating pleiotropy and annotation. *PLoS Genet.* 2014; 10(11):e1004787.doi: 10.1371/journal.pgen.1004787 [PubMed: 25393678]
- Cotsapas C, Voight BF, Rossin E, Lage K, Neale BM, Wallace C. ... Consortia FOno. Pervasive sharing of genetic effects in autoimmune disease. *PLoS Genet.* 2011; 7(8):e1002254.doi: 10.1371/journal.pgen.1002254 [PubMed: 21852963]
- Day FR, Loh PR, Scott RA, Ong KK, Perry JR. A Robust Example of Collider Bias in a Genetic Association Study. *Am J Hum Genet.* 2016; 98(2):392–393. DOI: 10.1016/j.ajhg.2015.12.019 [PubMed: 26849114]
- Ehret GB, Ferreira T, Chasman DI, Jackson AU, Schmidt EM, Johnson T, ... Munroe PB. The genetics of blood pressure regulation and its target organs from association studies in 342, 415 individuals. *Nat Genet.* 2016; doi: 10.1038/ng.3667
- Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, ... Johnson T. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature.* 2011; 478(7367):103–109. [PubMed: 21909115]
- Evangelou E, Warren HR, Mosen-Ansorena D, et al. Genetic analysis of over one million people identifies 535 novel loci for blood pressure. 2018 bioRxiv.
- Hackinger S, Zeggini E. Statistical methods to detect pleiotropy in human complex traits. *Open Biol.* 2017; 7(11)doi: 10.1098/rsob.170125
- Heid IM, Jackson AU, Randall JC, Winkler TW, Qi L, Steinthorsdottir V, ... Lindgren CM. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat Genet.* 2010; 42(11):949–960. DOI: 10.1038/ng.685 [PubMed: 20935629]
- Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet.* 2005; 6(2):95–108. DOI: 10.1038/nrg1521 [PubMed: 15716906]
- Hoffmann TJ, Ehret GB, Nandakumar P, Ranatunga D, Schaefer C, Kwok PY, ... Risch N. Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nat Genet.* 2017; 49(1):54–64. DOI: 10.1038/ng.3715 [PubMed: 27841878]
- Huang J, Johnson AD, O'Donnell CJ. PRIME: a method for characterization and evaluation of pleiotropic regions from multiple genome-wide association studies. *Bioinformatics.* 2011; 27(9):1201–1206. [PubMed: 21398673]
- Hunter DJ. Gene-environment interactions in human diseases. *Nat Rev Genet.* 2005; 6(4):287–298. DOI: 10.1038/nrg1578 [PubMed: 15803198]
- Ioannidis JP, Patsopoulos NA, Evangelou E. Heterogeneity in meta-analyses of genome-wide association investigations. *PLoS One.* 2007; 2(9):e841.doi: 10.1371/journal.pone.0000841 [PubMed: 17786212]
- Lane JM, Liang J, Vlasac I, Anderson SG, Bechtold DA, Bowden J, ... Saxena R. Genome-wide association analyses of sleep disturbance traits identify new loci and highlight shared genetics with neuropsychiatric and metabolic traits. *Nat Genet.* 2016; doi: 10.1038/ng.3749
- Liang J, Le TH, Edwards DRV, Tayo BO, Gaulton KJ, Smith JA, ... Franceschini N. Single-trait and multi-trait genome-wide association analyses identify novel loci for blood pressure in African-ancestry populations. *PLoS Genet.* 2017; 13(5):e1006728.doi: 10.1371/journal.pgen.1006728 [PubMed: 28498854]
- Liu C, Kraja AT, Smith JA, Brody JA, Franceschini N, Bis JC, ... Chasman DI. Meta-analysis identifies common and rare variants influencing blood pressure and overlapping with metabolic trait loci. *Nat Genet.* 2016; 48(10):1162–1170. DOI: 10.1038/ng.3660 [PubMed: 27618448]
- Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, ... Speliotes EK. Genetic studies of body mass index yield new insights for obesity biology. *Nature.* 2015; 518(7538):197–206. DOI: 10.1038/nature14177 [PubMed: 25673413]

- Marouli E, Graff M, Medina-Gomez C, Lo KS, Wood AR, Kjaer TR, ... Lettre G. Rare and low-frequency coding variants alter human adult height. *Nature*. 2017; 542(7640):186–190. DOI: 10.1038/nature21039 [PubMed: 28146470]
- Moonesinghe R, Khoury MJ, Liu T, Ioannidis JP. Required sample size and nonreplicability thresholds for heterogeneous genetic associations. *Proc Natl Acad Sci U S A*. 2008; 105(2):617–622. DOI: 10.1073/pnas.0705554105 [PubMed: 18174335]
- Nakaoka H, Inoue I. Meta-analysis of genetic association studies: methodologies, between-study heterogeneity and winner's curse. *J Hum Genet*. 2009; 54(11):615–623. DOI: 10.1038/jhg.2009.95 [PubMed: 19851339]
- Park H, Li X, Song YE, He KY, Zhu X. Multivariate Analysis of Anthropometric Traits Using Summary Statistics of Genome-Wide Association Studies from GIANT Consortium. *PLoS ONE*. 2016; 11(10):e0163912.doi: 10.1371/journal.pone.0163912 [PubMed: 27701450]
- Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet*. 2013; 14(7):483–495. DOI: 10.1038/nrg3461 [PubMed: 23752797]
- Turcot V, Lu Y, Highland HM, Schurmann C, Justice AE, Fine RS, ... Loos RJF. Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. *Nat Genet*. 2018; 50(1):26–41. DOI: 10.1038/s41588-017-0011-x [PubMed: 29273807]
- Turley P, Walters RK, Maghzian O, Okbay A, Lee JJ, Fontana MA. ... Social Science Genetic Association, C. Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat Genet*. 2018; 50(2):229–237. DOI: 10.1038/s41588-017-0009-4 [PubMed: 29292387]
- Vansteelandt S, Goetgeluk S, Lutz S, Waldman I, Lyon H, Schadt EE, ... Lange C. On the adjustment for covariates in genetic association analysis: a novel, simple principle to infer direct causal effects. *Genet Epidemiol*. 2009; 33(5):394–405. DOI: 10.1002/gepi.20393 [PubMed: 19219893]
- Warren HR, Evangelou E, Cabrera CP, Gao H, Ren M, Mifsud B. ... group UKBCCBw. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat Genet*. 2017; 49(3):403–415. DOI: 10.1038/ng.3768 [PubMed: 28135244]
- Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, ... Frayling TM. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet*. 2014; 46(11):1173–1186. DOI: 10.1038/ng.3097 [PubMed: 25282103]
- Zhu X, Feng T, Tayo BO, Liang J, Young JH, Franceschini N, ... Redline S. Meta-analysis of correlated traits via summary statistics from GWASs with an application in hypertension. *Am J Hum Genet*. 2015; 96(1):21–36. DOI: 10.1016/j.ajhg.2014.11.011 [PubMed: 25500260]

Figure 1.

Empirical power of tests for testing the association of a trait with a SNP at the 0.01 significance level under different sample sizes of the additional study for the covariate analysis when the SNP does not have any effect on the covariate. The sample sizes of the additional study (n_0) are 1000, 2000, 5,000 and 10,000 for each panel, respectively. The sample size of the study for the trait analysis (n_1) is 1,000. $\sigma_u = 2$. Noadj : regression without adjustment; Adj: regression with adjustment for the covariate. Q and Q_{hom} are indistinguishable.

Figure 2.

Empirical power of tests for testing the association of a trait with a SNP at the 0.01 significance level under different correlations between the traits and the covariate when the SNP does not have any effect on the covariate. $\sigma_u = 1, 2, 3$ and 4 for each panel, respectively. The sample size of the additional study for the covariate analysis (n_0) was 5,000. The sample size of the study for the trait analysis (n_1) is 1,000. Noadj: regression without adjustment; Adj: regression with adjustment for the covariate. Q and Q_{hom} are indistinguishable.

Figure 3.

Empirical power of tests for testing the association of a trait with a SNP at the 0.01 significance level under different sample sizes of the additional study for the covariate analysis when the SNP has pleiotropic effects on both the trait and the covariate. The sample sizes of the additional study (n_0) are 1000, 2000, 5,000 and 10,000 for each panel, respectively. The sample size of the study for the trait analysis (n_1) is 1,000. $\sigma_u = 2$. $b = 0.2$. Noadj : regression without adjustment; Adj: regression with adjustment for the covariate. Q and Q_{hom} are indistinguishable.

Figure 4.

Empirical power of tests for testing the association of a trait with a SNP at the 0.01 significance level under different correlations between the traits and the covariate when the SNP has pleiotropic effects on both the trait and the covariate. $\sigma_u = 1, 2, 3$ and 4 for each panel, respectively. $b = 0.2$. The sample size of the additional study for the covariate analysis (n_0) was 5,000. The sample size of the study for the trait analysis (n_1) is 1,000. Noadj: regression without adjustment; Adj: regression with adjustment for the covariate. Q and Q_{hom} are indistinguishable.

Figure 5.

Results for WC: (a) Venn Diagram of the significant SNPs ($P\text{-value} < 10^{-8}$) found by each of Q_{hom} , regression with adjustment for the covariate (Adj), and regression without adjustment for the covariate (Noadj). Q_{hom} , Adj and Noadj identified 780, 717 and 812 SNPs, respectively, associated with WC. (b) P-values of Q_{hom} plotted against P-values of Adj and P-values of Noadj.

Figure 6.

Results for WHR: (a) Venn Diagram of the significant SNPs (P-value < 10⁻⁸) found by each of Q_{hom} , regression with adjustment for the covariate (Adj), and regression without adjustment for the covariate (Noadj). Q_{hom} , Adj and Noadj identified 415, 620 and 399 SNPs, respectively, associated with WHR. (b) P-values of Q_{hom} plotted against P-values of Adj and P-values of Noadj.

Table 1

Empirical type I error under a homogeneous model with nominal type I error 0.01

n_0	b	σ_u	Adj_τ	Adj	Noadj	Q	Q_{hom}	Q_{het}
1000	0	0.5	0.010	0.010	0.010	0.010	0.010	0.010
		1	0.014	0.010	0.011	0.012	0.011	0.010
		1.5	0.027	0.010	0.011	0.012	0.011	0.010
		2	0.053	0.013	0.012	0.014	0.011	0.009
	0.1	0.5	0.009	0.009	0.009	0.009	0.009	0.009
		1	0.014	0.010	0.009	0.010	0.009	0.009
		1.5	0.027	0.018	0.011	0.012	0.010	0.008
		2	0.041	0.019	0.009	0.012	0.010	0.008
	0.2	0.5	0.010	0.012	0.009	0.009	0.009	0.009
		1	0.014	0.027	0.010	0.012	0.011	0.010
		1.5	0.025	0.041	0.009	0.011	0.009	0.008
		2	0.047	0.052	0.011	0.014	0.012	0.009
	0.3	0.5	0.012	0.017	0.012	0.012	0.012	0.012
		1	0.014	0.050	0.011	0.011	0.010	0.010
		1.5	0.025	0.097	0.010	0.011	0.010	0.008
		2	0.044	0.135	0.009	0.010	0.008	0.007
5000	0	0.5	0.012	0.011	0.011	0.012	0.012	0.012
		1	0.013	0.009	0.009	0.008	0.008	0.008
		1.5	0.026	0.011	0.010	0.013	0.012	0.011
		2	0.047	0.011	0.011	0.011	0.010	0.008
	0.1	0.5	0.010	0.010	0.010	0.009	0.010	0.010
		1	0.014	0.013	0.009	0.010	0.010	0.009
		1.5	0.027	0.019	0.012	0.013	0.013	0.011
		2	0.046	0.020	0.011	0.012	0.011	0.009
	0.2	0.5	0.012	0.013	0.011	0.011	0.011	0.012
		1	0.014	0.026	0.010	0.011	0.011	0.010
		1.5	0.025	0.044	0.010	0.011	0.011	0.010
		2	0.050	0.055	0.009	0.011	0.009	0.008
	0.3	0.5	0.011	0.017	0.011	0.011	0.011	0.011
		1	0.014	0.054	0.009	0.010	0.010	0.010
		1.5	0.026	0.100	0.010	0.011	0.011	0.010
		2	0.044	0.132	0.010	0.011	0.010	0.009

n_0 : the sample sizes of the covariate analysis; b : the SNP effect on the covariate; and σ_u : the shared effects on both the trait and the covariate.

Adj_τ: regression with adjustment for residual of the covariate; Adj: regression with adjustment for the covariate; Noadj: regression without adjustment; Q is the proposed method based on individual data; Q_{hom} is the proposed method based on the SNP effect estimated from a fixed effects meta-analysis; Q_{het} is the proposed method based on the BLUP estimate from a random effects meta-analysis.

The number of replicates for the simulations was 10,000.

Table 2

Empirical type I error under a heterogeneous model with nominal type I error 0.01

n_1	σ_b	σ_u	Noadj	Q	Q_{hom}	Q_{het}
1000	0.02	0.5	0.010	0.011	0.011	0.011
		1	0.009	0.011	0.010	0.010
		1.5	0.010	0.012	0.010	0.009
		2	0.011	0.013	0.010	0.008
	0.04	0.5	0.010	0.011	0.011	0.011
		1	0.010	0.012	0.011	0.010
		1.5	0.010	0.014	0.013	0.010
		2	0.011	0.013	0.011	0.009
	0.06	0.5	0.010	0.010	0.011	0.010
		1	0.008	0.011	0.011	0.008
		1.5	0.011	0.012	0.011	0.009
		2	0.010	0.014	0.011	0.008
	0.08	0.5	0.011	0.011	0.012	0.011
		1	0.010	0.013	0.012	0.010
		1.5	0.010	0.013	0.012	0.009
		2	0.007	0.012	0.010	0.006
	0.1	0.5	0.010	0.009	0.009	0.009
		1	0.010	0.013	0.012	0.010
		1.5	0.011	0.015	0.013	0.009
		2	0.010	0.017	0.014	0.009
5000	0.02	0.5	0.012	0.012	0.012	0.012
		1	0.010	0.011	0.010	0.009
		1.5	0.009	0.010	0.009	0.008
		2	0.009	0.012	0.011	0.009
	0.04	0.5	0.011	0.011	0.011	0.011
		1	0.010	0.011	0.011	0.010
		1.5	0.011	0.013	0.012	0.011
		2	0.009	0.012	0.011	0.009
	0.06	0.5	0.009	0.010	0.010	0.010
		1	0.011	0.012	0.012	0.011
		1.5	0.009	0.014	0.013	0.010
		2	0.009	0.016	0.015	0.011
	0.08	0.5	0.011	0.012	0.012	0.011
		1	0.010	0.015	0.014	0.011
		1.5	0.010	0.017	0.016	0.011
		2	0.009	0.019	0.018	0.012

n_1	σ_b	σ_u	Noadj	Q	Q _{hom}	Q _{het}
	0.1	0.5	0.009	0.010	0.010	0.009
		1	0.010	0.016	0.016	0.009
		1.5	0.008	0.021	0.020	0.011
		2	0.010	0.026	0.024	0.013

n0: the sample sizes of the covariate analysis; σ_b^2 : the heterogeneity level of the SNP effect on the covariate (i.e. the SNP effects in two groups were independently sampled from $N(b, \sigma_b^2)$, in which b was fixed at 0.2); σ_u : the shared effects on both the trait and the covariate.

Noadj: regression without adjustment; Q is the proposed method based on individual data; Qhom is the proposed method based on the SNP effect estimated from a fixed effects meta-analysis ; Qhet is the proposed method based on the BLUP estimate from a random effects meta-analysis.

The number of replicates for the simulations was 10,000.