ARTICLE

Mixed Model with Correction for Case-Control Ascertainment Increases Association Power

Tristan J. Hayeck,[1,2,*] Noah A. Zaitlen,[3] Po-Ru Loh,[2,4] Bjarni Vilhjalmsson,[2,4] Samuela Pollack,[2,4] Alexander Gusev,[2,4] Jian Yang,[5,6] Guo-Bo Chen,[5] Michael E. Goddard,[7] Peter M. Visscher,[5,6] Nick Patterson,[2] and Alkes L. Price[1,2,4,*]

We introduce a liability-threshold mixed linear model (LTMLM) association statistic for case-control studies and show that it has a well-controlled false-positive rate and more power than existing mixed-model methods for diseases with low prevalence. Existing mixed-model methods suffer a loss in power under case-control ascertainment, but no solution has been proposed. Here, we solve this problem by using a χ^2 score statistic computed from posterior mean liabilities (PMLs) under the liability-threshold model. Each individual's PML is conditional not only on that individual's case-control status but also on every individual's case-control status and the genetic relationship matrix (GRM) obtained from the data. The PMLs are estimated with a multivariate Gibbs sampler; the liability-scale phenotypic covariance matrix is based on the GRM, and a heritability parameter is estimated via Haseman-Elston regression on case-control phenotypes and then transformed to the liability scale. In simulations of unrelated individuals, the LTMLM statistic was correctly calibrated and achieved higher power than existing mixed-model methods for diseases with low prevalence, and the magnitude of the improvement depended on sample size and severity of case-control ascertainment. In a Wellcome Trust Case Control Consortium 2 multiple sclerosis dataset with >10,000 samples, LTMLM was correctly calibrated and attained a 4.3% improvement ($p = 0.005$) in χ^2 statistics over existing mixed-model methods at 75 known associated SNPs, consistent with simulations. Larger increases in power are expected at larger sample sizes. In conclusion, case-control studies of diseases with low prevalence can achieve power higher than that in existing mixed-model methods.

Introduction

Mixed-model association statistics are a widely used approach to correct for population structure and cryptic relatedness in genome-wide association studies (GWASs).[1–11] However, recent work shows that existing mixed-model association statistics have less power than standard logistic regression in studies of ascertained case and control subjects.[11] It is widely known that appropriate modeling of case-control ascertainment can produce substantial increases in power for case-control studies with fixed-effect covariates,[12–14] but such increases in power have not yet been achieved with models that include random effects.

We developed an association-score statistic based on a liability-threshold mixed linear model (LTMLM). The LTMLM statistic relies on the posterior mean liability (PML) of each individual; the PML is calculated with a multivariate Gibbs sampler.[15] The PML of each individual is conditional on the genetic relationship matrix (GRM), the case-control status of every individual, and the disease prevalence. Existing methods use a univariate prospective model to compute association statistics, but here we use a multivariate retrospective model.

The LTMLM statistic provides an increase in power in simulations of case-control studies of diseases with low prevalence, as demonstrated in settings of either simulated or real genotypes. In a Wellcome Trust Case Control Consortium 2 (WTCCC2) multiple sclerosis (MS) dataset with >10,000 samples, LTMLM was correctly calibrated and attained a 4.3% improvement ($p = 0.005$) in χ^2 statistics over existing mixed-model methods at 75 known associated SNPs, consistent with simulations.

Material and Methods

Overview of Method

We improve upon standard mixed-model methods[11] by using a retrospective association-score statistic (LTMLM) computed from PMLs under the liability-threshold model. The improvement over previous approaches comes from appropriate modeling of case-control ascertainment. We consider all individuals simultaneously and incorporate prevalence information.

Our method consists of three steps. First, the GRM is calculated, and a corresponding heritability parameter is estimated, modeling the phenotype covariance of all individuals (see "Estimation of Heritability Parameter"). The heritability parameter is estimated with Haseman-Elston (H-E) regression on the observed scale and then transformed to the liability scale. Second, PMLs are estimated with a truncated multivariate normal Gibbs sampler (see "PMLs"). The PML of each individual is conditional on that individual's case-control status, on every other individual's case-control status, and on disease prevalence and liability-scale phenotypic

[1]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA 02115, USA; [2]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA; [3]Lung Biology Center, School of Medicine, University of California, San Francisco, San Francisco, CA 94158,USA; [4]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA 02115, USA; [5]Queensland Brain Institute, University of Queensland, Brisbane, QLD 4072, Australia; [6]Diamantina Institute, University of Queensland, Brisbane, QLD 4072, Australia; [7]Faculty of Land and Food Resources, University of Melbourne, Melbourne, VIC 3010, Australia
*Correspondence: tjh488@mail.harvard.edu (T.J.H.), aprice@hsph.harvard.edu (A.L.P.)
http://dx.doi.org/10.1016/j.ajhg.2015.03.004. ©2015 by The American Society of Human Genetics. All rights reserved.

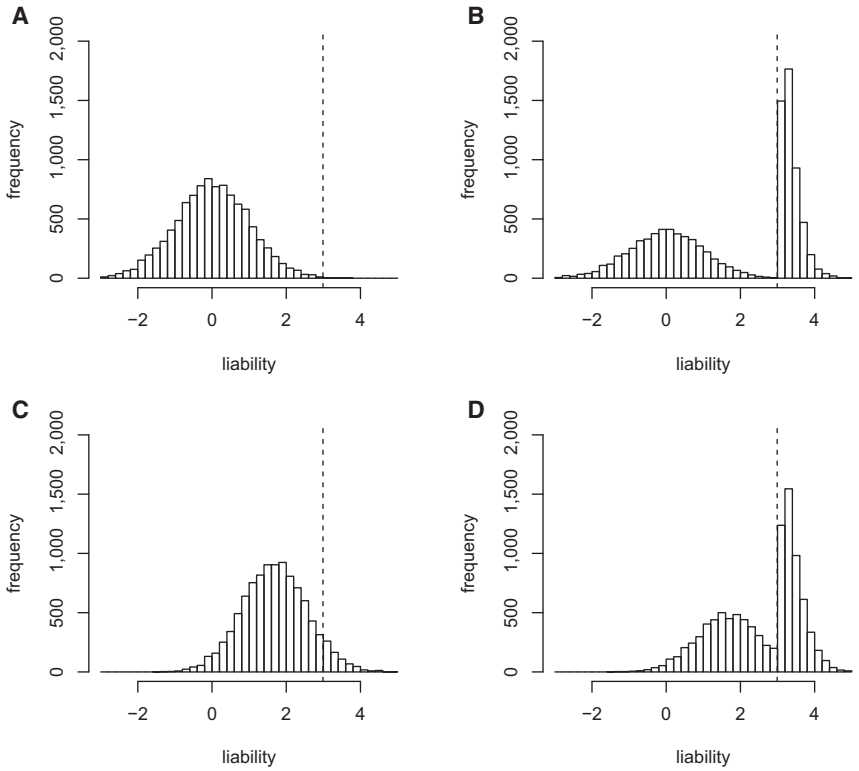

Figure 1. Genetic Relatedness to a Disease-Affected Case Subject Can Increase an Individual's PML

In (A) and (B), we plot distributions of liabilities for a set of 10,000 individuals under (A) random ascertainment or (B) case-control ascertainment for a disease with a prevalence of 0.1% (see Figure 2 in Lee et al.[16]). In (C) and (D), we plot the same distributions but condition on an individual's having a genetic relatedness of 0.5 with a disease-affected case subject while assuming a heritability of 1 on the liability scale.

covariance. Third, a χ^2 (1 degree of freedom) association-score statistic is computed on the basis of the association between the candidate SNP and the PML (see "LTMLM Association Statistic").

The toy example in Figure 1 illustrates how genetic relatedness to a disease-affected case subject can increase an individual's PML. In Figures 1A and 1B, we plot the distribution of liabilities in 10,000 unrelated individuals who were randomly ascertained and ascertained according to case-control status (for a disease with a prevalence of 0.1%), respectively. In Figures 1C and 1D, we plot the same distributions but condition on an individual's having a genetic relatedness of 0.5 with a disease-affected case subject while assuming a liability-scale heritability of 1.0. In each case, the posterior distribution of liabilities (and hence the PML) is shifted upward. (The magnitude and direction of this effect would be different for an individual with a genetic relatedness of 0.5 with a control subject). Our main focus below is on much lower levels of genetic relatedness (identity by state) among many unrelated samples, but the same principles apply.

Estimation of Heritability Parameter

Mixed-model association statistics rely on the estimation of a heritability parameter. We note that this heritability parameter, which Kang et al.[4] refer to as "pseudo-heritability," is generally lower than the total narrow-sense heritability (h^2) in datasets not dominated by family relatedness but can be larger than the heritability explained by genotyped SNPs (h_g^2)[17] in datasets with population structure or family relatedness. However, for ease of notation, we use the symbol h^2 to represent this heritability parameter. A list of all notation used below is provided in Table S1.

The goal is to test for association between a candidate SNP and a phenotype. We first consider a quantitative trait:

$$\boldsymbol{\varphi} = \boldsymbol{\beta x} + \boldsymbol{u} + \boldsymbol{e}. \qquad \text{(Equation 1)}$$

The phenotypic data (transformed to have mean 0 and variance 1) can be represented as a vector ($\boldsymbol{\varphi}$) with values for each individual (i). Genotype values of candidate SNPs are transformed to a vector ($\boldsymbol{x}$) with mean 0, variance 1, and effect size $\boldsymbol{\beta}$. The quantitative trait value depends on the fixed effect of the candidate SNP ($\boldsymbol{\beta x}$), the genetic random effect excluding the candidate SNP ($\boldsymbol{u}$), and the environmental component ($\boldsymbol{e}$). We extend to case-control traits via the liability-threshold model, in which each individual has an underlying, unobserved, normally distributed trait called the liability. An individual is affected by a disease if the liability exceeds a specified threshold (t), corresponding to disease prevalence[16] (Figure S1).

Standard mixed-model association methods generally estimate h^2 from a GRM and phenotypes by using restricted maximum likelihood (REML).[4,11] Genotypic data (excluding the candidate SNP[11]) are used for building a GRM:

$$\widehat{\boldsymbol{\Theta}} = \frac{\boldsymbol{X}^T\boldsymbol{X}}{M}, \qquad \text{(Equation 2)}$$

where $\boldsymbol{X}$ is a matrix of non-candidate SNPs normalized to mean 0 and variance 1 and M is the number of SNPs. We estimate h^2 by using H-E regression followed by a transformation to the liability scale. We obtain the H-E regression estimate by regressing the product of the case-control phenotypes on the off-diagonal terms of the GRM:[18,19]

$$\widehat{h^2_{\mathrm{HE}}} = \frac{\sum_{i\neq k}\boldsymbol{\pi}_i\boldsymbol{\pi}_k\widehat{\boldsymbol{\Theta}}_{ik}}{\sum_{i\neq k}\widehat{\boldsymbol{\Theta}}^2_{ik}}, \qquad \text{(Equation 3)}$$

where $\boldsymbol{\pi}_i$ denotes the case-control status of individual i and $\widehat{\boldsymbol{\Theta}}_{ik}$ is the genetic relatedness of individuals i and k. This gives an estimate on the observed scale, and then the estimate is transformed to the liability scale:[20]

$$\widehat{h^2_{\mathrm{HE,l}}} = \widehat{h^2_{\mathrm{HE}}}\frac{[K(1-K)]^2}{z^2(P(1-P))}, \qquad \text{(Equation 4)}$$

where z is the height of the standard normal density ($(1/2\pi)e^{-t^2/2}$) at the liability threshold t, K is disease prevalence, and P is the proportion of case subjects in the sample.[20]

Then, the variance between the individuals is modeled as the phenotypic covariance,

$$\boldsymbol{V} = h^2\widehat{\boldsymbol{\Theta}} + (\boldsymbol{I} - h^2)\boldsymbol{I}, \quad \text{(Equation 5)}$$

where $\widehat{\boldsymbol{\Theta}}$ is the $N \times N$ GRM, $\boldsymbol{V}$ is the phenotypic covariance, h^2 is the heritability parameter, and $\boldsymbol{I}$ is the identity matrix.

Using the phenotypic covariance matrix $\boldsymbol{V}$, the liability is modeled as a multivariate normal distribution:

$$L(\boldsymbol{\varphi}) = (2\pi)^{\frac{-n}{2}} |(\boldsymbol{V})|^{-1/2} \exp\left(\frac{-1}{2}(\boldsymbol{\varphi})^T(\boldsymbol{V})^{-1}(\boldsymbol{\varphi})\right). \quad \text{(Equation 6)}$$

We note that we observe the case-control phenotypes of the individuals and not the continuous liabilities.

PMLs

We first consider the univariate PML ($\mathbf{PML_{uni}}$), constructed independently for each individual; we generalize to the multivariate setting below. As described in Equations 11 and 12 in Lee et al.,[20] these correspond to the expected value of the liability and are conditional on the case-control status:

$$\mathrm{PML}_{\mathrm{uni,case}} = E[\boldsymbol{\varphi} \mid \boldsymbol{\pi}_i = 1] = z/$$

$$\mathrm{PML}_{\mathrm{uni,control}} = E[\boldsymbol{\varphi} \mid \boldsymbol{\pi}_i = 0] = -z/(1-K). \quad \text{(Equation 7)}$$

These values are calculated analytically in the univariate setting, and depending on case or control status, they can be thought of as the mean of a truncated normal distribution above or below, respectively, the liability threshold t.[20]

We next consider the multivariate PML ($\mathbf{PML_{multi}}$), estimated jointly across individuals. The $\mathbf{PML_{multi}}$ for each individual is conditional on that individual's case-control status, on every other individual's case-control status, and on their phenotypic covariance. The $\mathbf{PML_{multi}}$ is estimated with a Gibbs sampler, as in previous work[15] (which focused on family relatedness and did not consider association statistics). The Gibbs sampler is an iterative algorithm that generates random variables from conditional distributions in order to avoid the difficult task of explicitly calculating the marginal density for each random variable.

For each individual in turn, the conditional distribution of the liability is calculated on the basis of all of the other individuals, and a new value is generated. The algorithm is as follows:

Initialization: for each individual j,

$$\boldsymbol{\varphi}_i = \mathrm{PML}_{\mathrm{uni,case}} \text{ if } \boldsymbol{\pi}_i = 1 \text{ or } \boldsymbol{\varphi}_i = \mathrm{PML}_{\mathrm{uni,control}} \text{ if } \boldsymbol{\pi}_i = 0 \quad \text{(Equation 8)}$$

For each Markov Chain Monte Carlo (MCMC) iteration n

For each individual i
Sample $\boldsymbol{\varphi}_i$ from the constrained conditional univariate normal distribution
$L(\boldsymbol{\varphi}_i) \sim \exp(-\boldsymbol{\varphi}^T\mathbf{V}^{-1}\boldsymbol{\varphi}/2)$ and constraint $\boldsymbol{\varphi}_i \geq t$ if $\boldsymbol{\pi}_i = 1$, $\boldsymbol{\varphi}_i < t$ if $\boldsymbol{\pi}_i = 0$
(where $\boldsymbol{\varphi}_{\neq i}$ are fixed)

We use 100 burn-in iterations followed by 1,000 additional MCMC iterations. We estimate the $\mathbf{PML_{multi}}$ by averaging over MCMC iterations. We reduce the number of MCMC iterations needed via Rao-Blackwellization, which averages (across n iterations) the posterior means of the distributions from which each $\boldsymbol{\varphi}_i$ is sampled.

LTMLM Association Statistic

The LTMLM association statistic is calculated with the $\mathbf{PML_{multi}}$. For simplicity, we first consider the case where the liability is known. We jointly model the liability and the genotypes by using a retrospective model, enabling appropriate treatment of sample ascertainment. We concatenate the two vectors ($\boldsymbol{\varphi}$, $\boldsymbol{x}$) and derive the joint likelihood for these combined terms. The covariance of $\boldsymbol{\varphi}$ and $\boldsymbol{x}$ between individuals i and k is

$$\mathrm{Cov}(\boldsymbol{\varphi}_i, \boldsymbol{x}_k) = E[\boldsymbol{\varphi}_i, \boldsymbol{x}_k] - E[\boldsymbol{\varphi}_i]E[\boldsymbol{x}_k] = E[\boldsymbol{\varphi}_i, \boldsymbol{x}_k] = E[\boldsymbol{\beta}\boldsymbol{x}_i, \boldsymbol{x}_k] = \boldsymbol{\beta}\boldsymbol{\Theta}_{i,k}, \quad \text{(Equation 9)}$$

where $\boldsymbol{\Theta}$ is the true underlying GRM from which genotypes are sampled. (We note that $\boldsymbol{\Theta}$, which is unobserved, is different from the GRM $\widehat{\boldsymbol{\Theta}}$ estimated from the data). The variance of ($\boldsymbol{\varphi}$, $\boldsymbol{x}$) as a function of effect size $\boldsymbol{\beta}$ is

$$\boldsymbol{C}(\beta) = \begin{pmatrix} \boldsymbol{V} & \boldsymbol{\beta\Theta} \\ \boldsymbol{\beta\Theta}^T & \boldsymbol{\Theta} \end{pmatrix}. \quad \text{(Equation 10)}$$

Thus,

$$\boldsymbol{C}(\beta)^{-1} = \begin{pmatrix} \boldsymbol{V}^{-1} & -\boldsymbol{\beta}\boldsymbol{V}^{-1} \\ -\boldsymbol{\beta}(\boldsymbol{V}^{-1})^T & \boldsymbol{\Theta}^{-1} \end{pmatrix} + O(\beta^2), \quad \text{(Equation 11)}$$

where both of these matrices are $2N \times 2N$. (We note that the product of the matrices in Equations 10 and 11 is $\begin{pmatrix} \boldsymbol{I} + O(\beta^2) & 0 \\ 0 & \mathbf{I} + O(\beta^2) \end{pmatrix}$, whose difference from the identity contains only $O(\beta^2)$ terms.)

The joint likelihood of the liability and genotypes is distributed as a multivariate normal $N(0,\boldsymbol{C}(\beta))$, and thus

$$L(\boldsymbol{x}, \boldsymbol{\varphi} \mid \boldsymbol{\beta}) = (2\pi)^{\frac{-n}{2}} |\boldsymbol{C}(\beta)|^{-1/2} \exp\left(\frac{-1}{2}(\boldsymbol{\varphi}, \boldsymbol{x})^T\boldsymbol{C}(\beta)^{-1}(\boldsymbol{\varphi}, \boldsymbol{x})\right). \quad \text{(Equation 12)}$$

Taking the derivative of the log-likelihood results in the score equation. The determinant of the matrix $\boldsymbol{V}$ does not have any terms linear in $\boldsymbol{\beta}$, so the terms with $\boldsymbol{V}$ alone drop out when we take the derivative:

$$S(\boldsymbol{x}, \boldsymbol{\varphi} \mid \boldsymbol{\beta}) = \frac{d}{d\boldsymbol{\beta}} \ln L(\boldsymbol{x}, \boldsymbol{\varphi} \mid \boldsymbol{\beta}) = \frac{d}{d\boldsymbol{\beta}} \frac{-1}{2}(\boldsymbol{\varphi}, \boldsymbol{x})^T\boldsymbol{C}(\boldsymbol{\beta})(\boldsymbol{\varphi}, \boldsymbol{x})$$

$$= \frac{d}{d\boldsymbol{\beta}}(\boldsymbol{\varphi}, \boldsymbol{x})^T \begin{pmatrix} \boldsymbol{V}^{-1} & -\boldsymbol{\beta}\boldsymbol{V}^{-1} \\ -\boldsymbol{\beta}(\boldsymbol{V}^{-1})^T & \boldsymbol{\Theta}^{-1} \end{pmatrix} (\boldsymbol{\varphi}, \boldsymbol{x}) = \boldsymbol{V}^{-1}\boldsymbol{\varphi}\boldsymbol{x}. \quad \text{(Equation 13)}$$

The marginal score statistic tests the null hypothesis, which is that the fixed effect of the candidate SNP is 0 (H_0: $\boldsymbol{\beta} = 0$), against the alternative hypothesis (H_A: $\boldsymbol{\beta} \neq 0$). The denominator of the score statistic is the variance of the score evaluated under the null:

$$\mathrm{Var}(S(\boldsymbol{x}, \boldsymbol{\varphi} \mid \boldsymbol{\beta})) = (\boldsymbol{V}^{-1}\boldsymbol{\varphi})^T\boldsymbol{\Theta}(\boldsymbol{V}^{-1}\boldsymbol{\varphi}). \quad \text{(Equation 14)}$$

This leads to the score statistic

$$\frac{(\boldsymbol{x}^T\boldsymbol{V}^{-1}\boldsymbol{\varphi})^2}{(\boldsymbol{V}^{-1}\boldsymbol{\varphi})^T\boldsymbol{\Theta}(\boldsymbol{V}^{-1}\boldsymbol{\varphi})}, \quad \text{(Equation 15)}$$

where $\boldsymbol{\Theta}$, the true underlying genetic relatedness of the individuals, can be approximated by the identity matrix in datasets of unrelated individuals.

In Equations 9, 10, 11, 12, 13, 14, and 15, the liability was assumed to be known for simplicity. We now consider a case-control trait, with unobserved liability, and derive the score function by using the observed case-control status, π, of each individual.

Table 1. List of Association Statistics

	ATT	MLM	LTMLM
Quantitative or case-control trait	both	both	case-control
Estimates of heritability parameter	none	REML	H-E regression
Prospective or retrospective	prospective	prospective	retrospective
Equation	$\frac{(\boldsymbol{x}^T\pi^*)^2}{\boldsymbol{x}^T\boldsymbol{x}}$	$\frac{(\boldsymbol{x}^T\pi^*\boldsymbol{V}^{-1})^2}{\boldsymbol{x}^T\boldsymbol{V}^{-1}\boldsymbol{x}} = \frac{(\boldsymbol{x}^T\boldsymbol{V}^{-1}\mathbf{PML_{uni}})^2}{\boldsymbol{x}^T\boldsymbol{V}^{-1}\boldsymbol{x}}$	$\frac{(\boldsymbol{x}^T\boldsymbol{V}^{-1}\mathbf{PML_{multi}})^2}{(\boldsymbol{V}^{-1}\mathbf{PML_{multi}})^T\boldsymbol{I}(\boldsymbol{V}^{-1}\mathbf{PML_{multi}})}$
Corrects for confounding?	no	yes	yes
Models case-control ascertainment	no	no	yes

We list properties of the Armitage trend test (ATT), standard mixed-model association statistic (MLM), and proposed statistic (LTMLM). π^* is the normalized case-control status (mean 0 and variance 1), $\boldsymbol{x}$ is the normalized genotype, $\mathbf{PML_{uni}}$ is the univariate PML conditional on the case-control status of a single individual, $\mathbf{PML_{multi}}$ is the multivariate PML conditional of the case-control status of all individuals, $\boldsymbol{I}$ is the identity matrix, and $\boldsymbol{V}$ is the phenotypic covariance (on the observed scale for MLM and on the liability scale for LTMLM).

When we return to the score function and conditioning on case-control status,

$$S(\boldsymbol{x},\boldsymbol{\varphi}\,|\,\boldsymbol{\beta},\boldsymbol{\pi})_{\beta=0} = \frac{d}{d\boldsymbol{\beta}}\ln L(\boldsymbol{x},\boldsymbol{\varphi}\,|\,\boldsymbol{\beta},\ \boldsymbol{\pi})_{\beta=0} = \frac{\frac{dL(\boldsymbol{x},\boldsymbol{\varphi}\,|\,\boldsymbol{\beta},\pi)_{\beta=0}}{d\boldsymbol{\beta}}}{L(\boldsymbol{x},\boldsymbol{\varphi}\,|\,\boldsymbol{\beta},\pi)_{\beta=0}}. \quad \text{(Equation 16)}$$

When we introduce the unobserved quantitative liability, $\boldsymbol{\varphi}$, the score function can be rewritten in terms of the probability density of the liability:

$$\frac{\frac{dL(\boldsymbol{x},\boldsymbol{\varphi}\,|\,\boldsymbol{\beta})_{\beta=0}}{d\boldsymbol{\beta}}}{L(\boldsymbol{x},\boldsymbol{\varphi}\,|\,\boldsymbol{\beta})_{\beta=0}} = \frac{\int P(\boldsymbol{\varphi})\frac{dL(\boldsymbol{x},\boldsymbol{\varphi}\,|\,\boldsymbol{\beta})_{\beta=0}}{d\boldsymbol{\beta}}d\boldsymbol{\varphi}}{L(\boldsymbol{x},\boldsymbol{\varphi}\,|\,\boldsymbol{\beta})_{\beta=0}}$$

$$S(\boldsymbol{x},\boldsymbol{\varphi}\,|\,\boldsymbol{\beta},\pi) = \boldsymbol{C}\int P(\boldsymbol{\varphi})S(\boldsymbol{x},\boldsymbol{\varphi}\,|\,\boldsymbol{\beta},\boldsymbol{\pi})d\boldsymbol{\varphi} = S(\boldsymbol{x},E[\boldsymbol{\varphi}\,|\,\pi]\,|\,\boldsymbol{\beta}), \quad \text{(Equation 17)}$$

where $P(\boldsymbol{\varphi})$ is the probability density of the liability and $E[\boldsymbol{\varphi}|\ \pi]$ is the PML. It follows that an appropriate LTMLM score statistic is

$$\frac{(\boldsymbol{x}^T\boldsymbol{V}^{-1}\mathbf{PML_{multi}})^2}{(\boldsymbol{V}^{-1}\mathbf{PML_{multi}})^T\boldsymbol{\Theta}(\boldsymbol{V}^{-1}\mathbf{PML_{multi}})}. \quad \text{(Equation 18)}$$

Again, $\boldsymbol{\Theta}$ can be approximated by the identity matrix in datasets of unrelated individuals; we note that this choice affects only a constant calibration factor (because the denominator is the same for each candidate SNP) and that other calibration options are available (see below). As with other association statistics, the LTMLM score statistic generalizes to non-normally distributed genotypes.[21–23] The overall computational cost of computing the LTMLM statistic is $O(MN^2)$ when $M > N >$ number of iterations (Table S2). We have fixed the number of iterations at 100 burn-in iterations followed by 1,000 additional iterations.

We calculate the GRM via leave one chromosome out (LOCO) analysis, i.e., for each candidate SNP on a given chromosome, we calculate GRM by using all of the other chromosomes. This prevents deflation due to double counting of the candidate SNP as both a fixed effect and a random effect in the mixed model.[4,6,11]

Simulated Genotypes and Simulated Phenotypes

We performed simulations both by using simulated genotypes and simulated phenotypes and by using real genotypes and simulated phenotypes (see below). Quantitative liabilities for each individual were generated from SNP effects and an environmental component. All simulations included M candidate SNPs and an independent set of M SNPs used for calculating the GRM (so that candidate SNPs were not included in the GRM). For each scenario, a set of 100 simulations were run. We set ten candidate SNPs and ten GRM SNPs to be causal in simulations with $N = 1{,}000$ samples and set 50 candidate SNPs and 50 GRM SNPs to be causal in simulations with $N = 5{,}000$ samples to ensure that causal SNPs had similar average χ^2 statistics independent of M and N. We then dichotomized the resulting quantitative liabilities on the basis of the liability threshold to categorize each individual as a case or control subject. Case-control ascertainment was performed and simulated 50% case and 50% control subjects. We compared Armitage trend test (ATT), logistic regression (LogR), mixed linear model (MLM), and LTMLM statistics (see Table 1). MLM statistics were computed with the genome-wide complex trait analysis (GCTA)-LOCO statistic described in Yang et al.,[11] and the heritability parameters were estimated with the GCTA software.[24] We evaluated performance by using average χ^2 statistics at causal, null, and all markers, λ_{GC} at all markers (median χ^2 divided by 0.455),[25] and the proportion of causal and null markers that were significant at the p value thresholds of 0.05, 0.001, 1×10^{-6}, and 5×10^{-8}.

In the primary analyses, we simulated individuals without population structure or linkage disequilibrium (LD) with $N = 1{,}000$ or 5,000 samples; $M = 1{,}000$, 5,000, or 50,000 SNPs; and prevalence $K = 50\%$, 10%, 1%, or 0.1%. Genotypes were sampled from independent binomials with allele frequencies sampled from a uniform distribution over the values [0.1, 0.9]. In secondary analyses, we simulated population structure by simulating two populations with $F_{ST} = 0.01$ and allele frequencies drawn from beta distributions with parameters $p(1 - F_{ST})/F_{ST}$ and $(1 - p)(1 - F_{ST})/F_{ST}$, based on ancestral allele frequency p sampled from a uniform distribution over the values [0.1, 0.9].

To test the impact of the generative distribution, we simulated the underlying distribution by using a logit model instead of a liability-threshold model. The ten causal candidate SNPs were simulated with alternating fixed-effect sizes of $\boldsymbol{\beta} = 0.4$ or $\boldsymbol{\beta} = -0.4$. Then, case-control phenotypes were generated from a binomial distribution where the probability of being a case subject was $(\text{case}) = 1/(1 + \exp(-[c + \boldsymbol{\beta x}]))$, shifted by the affine term (c) and based on the desired disease prevalence.

WTCCC2 Genotypes and Simulated Phenotypes

We also conducted simulations by using real genotypes from WTCCC2 to incorporate LD and realistic population structure.

Table 2. Results for Simulated Genotypes and Simulated Phenotypes

Prevalence	Set	Statistic	ATT (SE)	LogR (SE)	MLM (SE)	LTMLM (SE)
***N* = 5,000 and *M* = 5,000**						
50%	causal	average	16.492 (0.325)	16.399 (0.321)	16.880 (0.332)	16.867 (0.332)
	null	average	0.988 (0.002)	0.988 (0.002)	0.990 (0.002)	0.989 (0.002)
	all	average	1.143 (0.004)	1.142 (0.004)	1.148 (0.004)	1.148 (0.004)
	all	λ_{GC}	1.010 (0.003)	1.010 (0.003)	1.014 (0.003)	1.014 (0.003)
25%	causal	average	18.637 (0.388)	18.509 (0.383)	19.014 (0.396)	19.056 (0.398)
	null	average	1.000 (0.002)	1.000 (0.002)	1.000 (0.002)	1.001 (0.002)
	all	average	1.177 (0.005)	1.175 (0.005)	1.180 (0.005)	1.181 (0.005)
	all	λ_{GC}	1.012 (0.004)	1.012 (0.004)	1.013 (0.004)	1.013 (0.004)
10%	causal	average	25.235 (0.501)	25.014 (0.492)	25.386 (0.506)	25.778 (0.514)
	null	average	0.993 (0.002)	0.992 (0.002)	0.991 (0.002)	0.992 (0.002)
	all	average	1.235 (0.006)	1.233 (0.006)	1.235 (0.006)	1.24 (0.007)
	all	λ_{GC}	1.005 (0.003)	1.005 (0.003)	1.007 (0.003)	1.008 (0.003)
1%	causal	average	45.376 (0.878)	44.682 (0.852)	42.594 (0.825)	46.691 (0.913)[a]
	null	average	1.000 (0.002)	0.999 (0.002)	0.990 (0.002)	1.000 (0.002)
	all	average	1.444 (0.011)	1.436 (0.011)	1.406 (0.010)	1.457 (0.011)
	all	λ_{GC}	1.020 (0.003)	1.020 (0.003)	1.011 (0.003)	1.019 (0.003)
0.1%	causal	average	68.648 (1.301)	67.099 (1.248)	56.303 (1.082)	70.810 (1.364)[a]
	null	average	1.000 (0.002)	1.000 (0.002)	0.918 (0.002)	1.000 (0.002)
	all	average	1.677 (0.016)	1.661 (0.016)	1.472 (0.013)	1.698 (0.017)
	all	λ_{GC}	1.026 (0.003)	1.026 (0.003)	0.942 (0.005)	1.025 (0.003)
***N* = 5,000 and *M* = 50,000**						
50%	causal	average	16.624 (0.331)	16.529 (0.327)	16.673 (0.332)	16.69 (0.333)
	null	average	1.000 (0.001)	0.999 (0.001)	1.000 (0.001)	1.000 (0.001)
	all	average	1.015 (0.001)	1.015 (0.001)	1.016 (0.001)	1.015 (0.001)
	all	λ_{GC}	1.003 (0.001)	1.002 (0.001)	0.999 (0.001)	0.999 (0.001)
25%	causal	average	18.965 (0.37)	18.843 (0.366)	19.030 (0.372)	19.040 (0.372)
	null	average	1.003 (0.001)	1.003 (0.001)	1.003 (0.001)	1.003 (0.001)
	all	average	1.021 (0.001)	1.020 (0.001)	1.021 (0.001)	1.021 (0.001)
	all	λ_{GC}	1.006 (0.001)	1.006 (0.001)	1.003 (0.001)	1.003 (0.001)
10%	causal	average	23.710 (0.444)	23.528 (0.437)	23.868 (0.449)	23.910 (0.450)
	null	average	1.001 (0.001)	1.000 (0.001)	1.001 (0.001)	1.001 (0.001)
	all	average	1.023 (0.001)	1.023 (0.001)	1.024 (0.001)	1.024 (0.001)
	all	λ_{GC}	1.007 (0.001)	1.007 (0.001)	1.005 (0.001)	1.004 (0.001)
1%	causal	average	46.683 (0.883)	45.969 (0.859)	46.44 (0.881)	47.368 (0.905)
	null	average	0.999 (0.001)	0.999 (0.001)	0.999 (0.001)	0.999 (0.001)
	all	average	1.045 (0.001)	1.044 (0.001)	1.045 (0.001)	1.045 (0.001)
	all	λ_{GC}	1.004 (0.001)	1.004 (0.001)	1.001 (0.001)	1.000 (0.001)

(*Continued on next page*)

Table 2. *Continued*

Prevalence	Set	Statistic	ATT (SE)	LogR (SE)	MLM (SE)	LTMLM (SE)
0.1%	causal	average	67.059 (1.278)	65.561 (1.225)	65.232 (1.251)	68.618 (1.333)[a]
	null	average	0.999 (0.001)	0.999 (0.001)	0.999 (0.001)	0.999 (0.001)
	all	average	1.065 (0.002)	1.063 (0.002)	1.063 (0.002)	1.067 (0.002)
	all	λ_{GC}	1.004 (0.001)	1.004 (0.001)	1.000 (0.001)	1.000 (0.001)

We report average χ^2 statistics across 100 simulations for each individual scenario. *N* is the number of individuals, and *M* is the number of SNPs. The set indicates all SNPs, the 1% causal SNPs, or the 99% null SNPs. The disease prevalence ranges from 50% (no case-control ascertainment) to 0.1%.
[a]The settings where LTMLM demonstrated at least a 5% improvement over MLM in χ^2 statistics at causal markers.

The WTCCC2 data contained 360,557 SNPs and 15,633 samples, as described previously.[11] Because the goal of the power study is to demonstrate a comparison of the statistics under case-control ascertainment, we used $N = 1{,}000$ samples (500 case and 500 control subjects) and simulated phenotypes with a prevalence of 50%, 25%, and 10%. The prevalence was restricted to a lower bound of 10% because of the limitation of having only 15,633 WTCCC2 samples for simulating case-control ascertainment. We computed ATT, LogR, MLM, and LTMLM statistics as described above.

WTCCC2 Genotypes and MS Phenotypes

Finally, we analyzed WTCCC2 individuals with ascertained case-control phenotypes for MS,[11] a disease with a prevalence of around 0.1%. As in previous work, we assume that the disease prevalence is known on the basis of external epidemiological literature.[20,26,27] For the WTCCC2 MS data, we used a threshold of 3.0, corresponding to a disease prevalence of 0.1%.[26] We computed ATT, LogR, MLM, and LTMLM statistics as described above. Although the underlying MS study was appropriately matched for ancestry,[28] the data made available to researchers included only pan-European case and UK control subjects. Thus, the WTCCC2 dataset shows a severe mismatch in ancestry of case and control subjects; this severe mismatch between case and control subjects is not representative of a typical GWAS. We thus restricted our primary analysis to 10,034 samples with only a moderate mismatch in ancestry, but we also performed analyses of unmatched and stringently matched datasets (Figure S2). The unmatched dataset contained 10,204 case and 5,429 control subjects. We performed matching by first calculating 20 principal components (PCs) in the full cohort and weighing the contribution of each PC on the basis of how much phenotype variance it explained in a multiple regression. A Euclidean distance over these 20 weighted dimensions was then computed for all pairs of individuals, and each case subject was greedily assigned the nearest unmatched control subject until no matched case-control pairs could be identified. Finally, any matched case-control pairs who were not within 6 SDs of the mean pairwise distance were removed as outliers, yielding the 5,017 case and 5,017 matched control subjects used in our primary analysis. We performed stringent matching by additionally removing any matched case-control pairs who were not within 2 SDs of the mean pairwise distance, yielding the 4,094 case and 4,094 matched control subjects used in our stringently matched analysis.

We compared association statistics at 75 published SNPs associated with MS.[11] We used a jackknife approach to assess the statistical significance of differences in association statistics by excluding each of the 75 published SNPs in turn.

Results

Simulations: Simulated Genotypes and Simulated Phenotypes

We first conducted simulations by using simulated genotypes and simulated ascertained case-control phenotypes (see Material and Methods). Our main simulations involved unrelated individuals with no population structure, but the impact of population structure is explored below. We evaluated the power of ATT, LogR, MLM, and LTMLM. We report average χ^2 statistics at causal, null, and all markers and λ_{GC} at all markers in Table 2, and we report the proportion of causal and null markers that were significant at various p value thresholds in Table S3. For diseases with low prevalence, the LTMLM statistic outperformed the ATT, LogR, and MLM statistics. Improvements in average χ^2 statistics at causal markers (which are naturally interpreted as the increase in effective sample size) were larger than improvements in power to detect an association at a given p value threshold, most likely as a result of the variable (normally distributed) effect sizes in this simulation (see below). For LTMLM and MLM at disease prevalences of 0.1%, 26%, and 5%, improvements in average χ^2 statistics at causal markers were observed in simulations with 5,000 SNPs and 50,000 SNPs, respectively. Smaller improvements were observed at higher disease prevalences. Test statistics were well calibrated at null markers. Simulations at other values of *M* and *N* indicated that the magnitude of the improvement depends on the value of N/M (Tables S3 and S4). Simulations with population structure demonstrated similar results but also showed inflation in the ATT statistic as expected (Tables S5 and S6).

The MLM statistics were calculated with an h^2 parameter estimated via REML,[4] but the LTMLM statistics were calculated with an h^2 parameter first estimated via H-E regression on case-control phenotypes and then transformed to the liability scale[18,20] (see Material and Methods). As case-control ascertainment became more severe, the H-E regression estimate of h^2 remained unbiased, whereas the variance-component estimate was severely downwardly biased even after transformation to the liability scale (Table 3 and Table S7), consistent with previous work (see Golan et al.[29] and Table S9 in Yang et al.[11]). Population structure resulted in bias for both REML and H-E-regression

Table 3. Heritability-Parameter Estimates from Simulated Genotypes and Phenotypes

	Liability Scale		Observed Scale	
Prevalence	**H-E (SE)**	**REML (SE)**	**H-E (SE)**	**REML (SE)**
***N* = 5,000 and *M* = 5,000**				
50%	0.255 (0.006)	0.253 (0.006)	0.162 (0.004)	0.161 (0.004)
25%	0.24 (0.005)	0.236 (0.004)	0.172 (0.003)	0.17 (0.003)
10%	0.24 (0.004)	0.23 (0.004)	0.228 (0.004)	0.219 (0.004)
1%	0.25 (0.004)	0.21 (0.002)	0.453 (0.007)	0.381 (0.004)
0.1%	0.253 (0.003)	0.158 (0.000)	0.719 (0.009)	0.449 (0.001)
***N* = 5,000 and *M* = 50,000**				
50%	0.272 (0.009)	0.274 (0.009)	0.173 (0.006)	0.174 (0.006)
25%	0.246 (0.010)	0.252 (0.010)	0.177 (0.007)	0.181 (0.007)
10%	0.225 (0.004)	0.231 (0.004)	0.214 (0.004)	0.219 (0.004)
1%	0.246 (0.004)	0.241 (0.004)	0.445 (0.008)	0.437 (0.008)
0.1%	0.259 (0.004)	0.241 (0.004)	0.734 (0.012)	0.684 (0.01)

We report results, on both the liability and observed scales, averaged over 100 simulations. The true h^2, explained by the SNPs used for building the GRM, is 25% on the liability scale for all simulations. These results correspond to the same sets of simulations in Table 2.

estimates of h^2, although the bias for the REML estimates was higher (Table S8). These biases did not inflate LTMLM or MLM statistics under the null hypothesis (Tables S5 and S6). We note that previous work has shown that running MLM with the correct h^2 parameter does not ameliorate the loss in power for MLM.[11]

We also evaluated performance in settings where the liability threshold (equivalently, the disease prevalence) was misspecified (Tables S9 and S10). The LTMLM statistic remained properly calibrated under the null and continued to outperform the MLM statistic given that the impact of misspecifying the liability threshold was small. Misspecifying the liability threshold led to bias in liability-scale heritability estimates as a result of the inaccurate conversion from the observed scale to the liability scale (Table S11).

Finally, we evaluated performance when phenotypes were generated with a logit model instead of a liability-threshold model; for simplicity, we used a fixed effect size for causal candidate SNPs (see Material and Methods). At low disease prevalence, we observed improvements for LTMLM both in average χ^2 statistics at causal markers (Table S12) and in power to detect an association at a given p value threshold (Table S13); improvements in power depend heavily on the distribution of causal effect sizes and are larger in simulations with fixed causal effect sizes than in simulations with variable causal effect sizes.

Simulations: WTCCC2 Genotypes and Simulated Phenotypes

We next conducted simulations by using real WTCCC2 genotypes and simulated phenotypes from ascertained case and control individuals (see Material and Methods).[11,28] For a given value of M (M SNPs for calculating the GRM plus M candidate SNPs for a total of $2M$ SNPs), we used the first $M/2$ SNPs from each of the first four chromosomes. The GRM was calculated with SNPs on chromosomes 3 and 4, and SNPs on chromosomes 1 and 2 were treated as the candidate SNPs. The simulated phenotypes were generated from chromosomes 1 and 3, where 1% of the SNPs were randomly selected as being causal. Results are reported for causal SNPs on chromosome 1 and for null SNPs on chromosome 2, the latter of which were not used for building the GRM.

Results for simulations including 1,000 and 10,000 SNPs (M) and sample sizes fixed at 500 case and 500 control subjects are displayed in Table 4 and Tables S14 and S15. Once again, the LTMLM statistic outperformed ATT and MLM as case-control ascertainment became more severe. (A limitation of these simulations is that performing case-control ascertainment on a fixed set of individuals limits case-control sample size; thus, these simulations were restricted to a disease prevalence of 10% or higher. It is reasonable to infer that for rarer diseases with more extreme case-control ascertainment the LTMLM statistic would achieve even higher power gains, as was demonstrated in simulations with simulated genotypes.)

The h^2 parameter estimates for simulations with real genotypes are displayed in Table 5. The H-E regression estimates are unbiased, but the REML estimates are again downwardly biased at lower prevalence and large N/M.

WTCCC2 MS Dataset

We analyzed the WTCCC2 genotypes together with MS case-control phenotypes: 5,172 MS case subjects and 5,172 control subjects were genotyped on Illumina chips[11,28](see Material and Methods). We compared ATT, ATT with five PCs (ATT + PCs),[22] LogR, LogR with five PCs (LogR + PCs), MLM, and LTMLM. We evaluated calibration by using the average χ^2 and λ_{GC} over all SNPs; we note that the average χ^2 and λ_{GC} are expected to be greater than 1 as a result of polygenic effects.[11,30] We believe that LTMLM is effective in correcting for confounding and that a higher value of λ_{GC} for LTMLM than for MLM is most likely due to true polygenic signal and reflects the higher power of LTMLM.

We evaluated power by using the average χ^2 over the 75 published SNPs (Table 6) and the proportion of published SNPs that were significant at various p value thresholds (Table S16). The LTMLM method performed best: it had a 4.3% improvement in average χ^2 statistics scaled by λ_{GC} over MLM (jackknife $p = 0.005$; see Material and Methods) and an even larger improvement over ATT and ATT + PCs, consistent with simulations (Table 2 and Table S3). LTMLM also detected 56/75 known associations as nominally significant ($p < 0.05$) after λ_{GC} correction, whereas MLM detected only 53/75, although this difference is not statistically significant. Similar results were obtained when association statistics were calibrated via LD-score regression[31]

Table 4. Results on Real Genotypes and Simulated Phenotypes

Prevalence	Set	Statistic	ATT (SE)	ATT + PCs (SE)	LogR (SE)	LogR + PCs (SE)	MLM (SE)	LTMLM (SE)
***M* = 1,000**								
50%	causal	average	16.234 (0.723)	15.775 (0.705)	15.566 (0.674)	15.14 (0.658)	17.48 (0.780)	17.412 (0.775)
	null	average	1.444 (0.011)	1.425 (0.010)	1.433 (0.010)	1.415 (0.010)	1.472 (0.011)	1.471 (0.011)
	all	average	1.592 (0.014)	1.569 (0.013)	1.574 (0.013)	1.552 (0.013)	1.632 (0.015)	1.630 (0.015)
	all	λ_{GC}	1.214 (0.017)	1.224 (0.017)	1.214 (0.017)	1.223 (0.017)	1.234 (0.017)	1.226 (0.016)
25%	causal	average	19.277 (0.771)	18.581 (0.754)	18.47 (0.721)	17.822 (0.706)	20.493 (0.820)	20.642 (0.831)
	null	average	1.551 (0.012)	1.507 (0.011)	1.538 (0.012)	1.495 (0.011)	1.571 (0.012)	1.577 (0.013)
	all	average	1.728 (0.015)	1.678 (0.014)	1.707 (0.015)	1.658 (0.014)	1.761 (0.016)	1.768 (0.016)
	all	λ_{GC}	1.256 (0.016)	1.245 (0.017)	1.255 (0.016)	1.245 (0.017)	1.241 (0.015)	1.243 (0.015)
10%	causal	average	22.838 (0.961)	21.406 (0.878)	21.618 (0.881)	20.336 (0.808)	23.865 (1.022)	24.661 (1.064)
	null	average	1.664 (0.014)	1.583 (0.012)	1.647 (0.014)	1.568 (0.012)	1.668 (0.014)	1.695 (0.015)
	all	average	1.876 (0.018)	1.781 (0.016)	1.846 (0.017)	1.756 (0.015)	1.890 (0.019)	1.925 (0.019)
	all	λ_{GC}	1.271 (0.016)	1.285 (0.015)	1.270 (0.016)	1.284 (0.015)	1.251 (0.017)	1.270 (0.016)
***M* = 10,000**								
50%	causal	average	16.898 (0.702)	16.815 (0.698)	16.229 (0.659)	16.156 (0.655)	17.279 (0.725)	17.26 (0.725)
	null	average	1.078 (0.002)	1.083 (0.002)	1.074 (0.002)	1.078 (0.002)	1.078 (0.002)	1.076 (0.002)
	all	average	1.094 (0.002)	1.099 (0.002)	1.089 (0.002)	1.094 (0.002)	1.095 (0.002)	1.092 (0.002)
	all	λ_{GC}	1.039 (0.005)	1.047 (0.005)	1.039 (0.005)	1.046 (0.005)	1.039 (0.005)	1.035 (0.005)
25%	causal	average	17.573 (0.726)	17.293 (0.696)	16.856 (0.679)	16.61 (0.654)	17.976 (0.749)	18.077 (0.758)
	null	average	1.076 (0.002)	1.078 (0.002)	1.071 (0.002)	1.073 (0.002)	1.074 (0.002)	1.073 (0.002)
	all	average	1.092 (0.002)	1.094 (0.002)	1.087 (0.002)	1.089 (0.002)	1.091 (0.002)	1.090 (0.002)
	all	λ_{GC}	1.040 (0.004)	1.045 (0.004)	1.040 (0.004)	1.044 (0.004)	1.040 (0.004)	1.039 (0.004)
10%	causal	average	24.379 (1.026)	24.116 (1.014)	23.127 (0.944)	22.894 (0.934)	24.987 (1.071)	25.399 (1.091)
	null	average	1.112 (0.002)	1.115 (0.002)	1.107 (0.002)	1.110 (0.002)	1.108 (0.002)	1.111 (0.002)
	all	average	1.136 (0.003)	1.138 (0.003)	1.129 (0.002)	1.132 (0.002)	1.131 (0.003)	1.135 (0.003)
	all	λ_{GC}	1.051 (0.005)	1.059 (0.004)	1.050 (0.005)	1.058 (0.004)	1.044 (0.004)	1.047 (0.004)

We report average χ^2 statistics. *M* is the number of SNPs, and sample size is fixed at 500 case and 500 control subjects.

(Table S17). A perfectly matched dataset with 4,094 MS case subjects and 4,094 control subjects yielded a similar improvement for LTMLM over MLM (Table S18). We also applied LTMLM to the full unmatched dataset of 10,204 MS subjects and 5,429 control subjects, where a severe mismatch in ancestry between case and control subjects was not representative of a typical GWAS. The LOCO estimates of h^2 demonstrated inflation before we controlled for population structure (Table S19). In this analysis, the H-E regression estimate of h^2 produced an unrealistic value of 7.3 on the observed scale (corresponding to 2.8 on the liability scale), which is outside the plausible 0–1 range, suggesting severe population stratification or other severe problems with the data. We do not recommend the use of LTMLM on unmatched samples when such severe problems are detected. For completeness, we report the results of running LTMLM, which results in a loss of power (Table S18).

Discussion

We have shown that controlling for case-control ascertainment by using the LTMLM statistic can lead to significant power improvements in case-control studies of diseases of low prevalence. This was demonstrated via simulations using both simulated and real genotypes and in WTCCC2 MS case-control data. We emphasize that the improvement applies to case-control studies of diseases with low prevalence. We note that logistic and linear regression generally produce similar results and that logistic mixed-model score tests that do not explicitly model case-control ascertainment are likely to produce results similar to those of standard linear mixed-model methods.

The LTMLM statistic should not be used if the inferred liability-scale h^2 parameter is outside the plausible 0–1 bound, given that this is indicative of severe population stratification or other severe problems with the data (this

Table 5. Heritability-Parameter Estimates on Real Genotypes and Simulated Phenotypes

	Liability Scale		Observed Scale	
Prevalence	**H-E (SE)**	**REML (SE)**	**H-E (SE)**	**REML (SE)**
***M* = 10,000**				
50%	0.259 (0.013)	0.252 (0.010)	0.165 (0.008)	0.161 (0.006)
25%	0.241 (0.010)	0.238 (0.008)	0.173 (0.007)	0.171 (0.006)
10%	0.245 (0.011)	0.242 (0.007)	0.233 (0.010)	0.230 (0.007)
***M* = 10,000**				
50%	0.236 (0.014)	0.245 (0.013)	0.150 (0.009)	0.156 (0.008)
25%	0.250 (0.014)	0.264 (0.013)	0.180 (0.010)	0.190 (0.010)
10%	0.259 (0.012)	0.261 (0.009)	0.246 (0.011)	0.248 (0.009)

These results are from the same simulations used to generate Table 4 and Tables S14 and S15. We report results on both the liability and the observed scales. The true h^2 explained by the SNPs used for building the GRM is 25% on the liability scale for all simulations.

can also be assessed via PC analysis; see Figure S2). In such settings, either matching based on ancestry should first be performed or other statistics should be used.

Several limitations of LTMLM are directions for future study. First, previous work has shown that using the PMLs in conjunction with fixed effects such as body mass index, age, or known associated SNPs will further increase power.[12,20] The incorporation of fixed-effect covariates into the LTMLM statistic is not considered here and remains a future direction. Second, the calibration of our statistic in unrelated samples relies on an approximation that works well in the WTCCC2 data analyzed but that might not work well in all datasets. Here, calibration via LD-score regression offers an appealing alternative.[31] Third, we did not consider case-control studies in family datasets, which also represents a future direction. The LTMLM score statistic in its current form is appropriate for association testing in population case-control samples with low levels of relatedness. In family datasets, other approaches to calibration, such as LD-score regression, could potentially be explored.[31] Fourth, the method relies on the assumption of an underlying normally distributed liability. Although this assumption is widely accepted by many geneticists,[32,33] and the method also performs well under a different generative model (Tables S12 and S13), further work on whether case-control traits are accurately modeled by normally distributed liabilities is warranted. Fifth, the method does not estimate odds ratios; in this respect, the method is similar to other mixed-model association methods.[4,5,8,11] However, liability-scale effect sizes can be converted to odds ratios.[13] Sixth, analogous to standard mixed-model association methods, LTMLM requires running time $O(MN^2)$ when $M > N >$ number of MCMC iterations. This might be computationally intractable in very large datasets. We are developing much faster mixed-model methods,[34] but these methods do not consider case-control ascertainment and should not be applied to ascertained case-control data for diseases of low prevalence. The incorporation of the ideas we have described here into these methods is an open question. Seventh, potential application of the LTMLM statistic to rare-variant datasets is not considered here and remains a future direction. Finally, our methods could potentially be extended to multiple traits.[7,27,35]

Supplemental Data

Supplemental Data include two figures and 19 tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2015.03.004.

Acknowledgments

We thank D. Golan and S. Rosset for helpful discussions. This research was funded by NIH grants R01 HG006399 and T32 GM074897. This study made use of data generated by the Wellcome Trust Case Control Consortium. A full list of the investigators who contributed to the generation of the data is available at http://www.wtccc.org.uk/.

Received: August 21, 2014
Accepted: March 5, 2015
Published: April 16, 2015

Web Resources

The URLs for data presented herein are as follows:

GCTA (Genome-wide Complex Trait Analysis), http://www.complextraitgenomics.com/software/gcta/

Liability threshold mixed linear model (LTMLM) in LTSOFT, http://www.hsph.harvard.edu/alkes-price/software/

Table 6. Results on WTCCC2 MS Dataset

Category	Metric	ATT (SE)	ATT + PCs (SE)	LogR (SE)	LogR + PCs (SE)	MLM (SE)	LTMLM (SE)
Published	mean	11.661 (1.169)	9.98 (0.984)	11.619 (1.161)	9.871 (0.965)	9.919 (0.974)	10.587 (1.017)
Genome-wide	mean	1.379 (0.003)	1.152 (0.003)	1.378 (0.003)	1.142 (0.003)	1.144 (0.003)	1.172 (0.003)
Genome-wide	λ_{GC}	1.343	1.125	1.343	1.115	1.115	1.141
Published	mean/λ_{GC}	8.695	8.880	8.663	8.860	8.905	9.292

We report, going down the rows, the average χ^2 over 75 published SNPs, the average χ^2 across the entire genome, including 360,557 SNPs, the genome-wide λ_{GC}, and then the average χ^2 across 75 published SNPs after normalization by the λ_{GC}. All results are based on analysis of 10,034 individuals (see main text).

References

1. Yu, J., Pressoir, G., Briggs, W.H., Vroh Bi, I., Yamasaki, M., Doebley, J.F., McMullen, M.D., Gaut, B.S., Nielsen, D.M., Holland, J.B., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. *38*, 203–208.
2. Zhao, K., Aranzana, M.J., Kim, S., Lister, C., Shindo, C., Tang, C., Toomajian, C., Zheng, H., Dean, C., Marjoram, P., and Nordborg, M. (2007). An Arabidopsis example of association mapping in structured samples. PLoS Genet. *3*, e4.
3. Kang, H.M., Zaitlen, N.A., Wade, C.M., Kirby, A., Heckerman, D., Daly, M.J., and Eskin, E. (2008). Efficient control of population structure in model organism association mapping. Genetics *178*, 1709–1723.
4. Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S.Y., Freimer, N.B., Sabatti, C., and Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. Nat. Genet. *42*, 348–354.
5. Svishcheva, G.R., Axenovich, T.I., Belonogova, N.M., van Duijn, C.M., and Aulchenko, Y.S. (2012). Rapid variance components-based method for whole-genome association analysis. Nat. Genet. *44*, 1166–1170.
6. Listgarten, J., Lippert, C., Kadie, C.M., Davidson, R.I., Eskin, E., and Heckerman, D. (2012). Improved linear mixed models for genome-wide association studies. Nat. Methods *9*, 525–526.
7. Korte, A., Vilhjálmsson, B.J., Segura, V., Platt, A., Long, Q., and Nordborg, M. (2012). A mixed-model approach for genome-wide association studies of correlated traits in structured populations. Nat. Genet. *44*, 1066–1071.
8. Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. *44*, 821–824.
9. Segura, V., Vilhjálmsson, B.J., Platt, A., Korte, A., Seren, Ü., Long, Q., and Nordborg, M. (2012). An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat. Genet. *44*, 825–830.
10. Listgarten, J., Lippert, C., and Heckerman, D. (2013). FaST-LMM-Select for addressing confounding from spatial structure and rare variants. Nat. Genet. *45*, 470–471.
11. Yang, J., Zaitlen, N.A., Goddard, M.E., Visscher, P.M., and Price, A.L. (2014). Advantages and pitfalls in the application of mixed-model association methods. Nat. Genet. *46*, 100–106.
12. Zaitlen, N., Lindström, S., Pasaniuc, B., Cornelis, M., Genovese, G., Pollack, S., Barton, A., Bickeböller, H., Bowden, D.W., Eyre, S., et al. (2012). Informed conditioning on clinical covariates increases power in case-control association studies. PLoS Genet. *8*, e1003032.
13. Zaitlen, N., Pasaniuc, B., Patterson, N., Pollack, S., Voight, B., Groop, L., Altshuler, D., Henderson, B.E., Kolonel, L.N., Le Marchand, L., et al. (2012). Analysis of case-control association studies with known risk variants. Bioinformatics *28*, 1729–1737.
14. Mefford, J., and Witte, J.S. (2012). The Covariate's Dilemma. PLoS Genet. *8*, e1003096.
15. Campbell, D.D., Sham, P.C., Knight, J., Wickham, H., and Landau, S. (2010). Software for generating liability distributions for pedigrees conditional on their observed disease states and covariates. Genet. Epidemiol. *34*, 159–170.
16. Falconer, D.S. (1967). The inheritance of liability to diseases with variable age of onset, with particular reference to diabetes mellitus. Ann. Hum. Genet. *31*, 1–20.
17. Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. Nat. Genet. *42*, 565–569.
18. Haseman, J.K., and Elston, R.C. (1972). The investigation of linkage between a quantitative trait and a marker locus. Behav. Genet. *2*, 3–19.
19. Chen, G.-B. (2014). Estimating heritability of complex traits from genome-wide association studies using IBS-based Haseman-Elston regression. Front. Genet. *5*, 107.
20. Lee, S.H., Wray, N.R., Goddard, M.E., and Visscher, P.M. (2011). Estimating missing heritability for disease from genome-wide association studies. Am. J. Hum. Genet. *88*, 294–305.
21. Thornton, T., and McPeek, M.S. (2007). Case-control association testing with related individuals: a more powerful quasi-likelihood score test. Am. J. Hum. Genet. *81*, 321–337.
22. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. *38*, 904–909.
23. Heyde, C.C. (1997). Quasi-likelihood and its application: a general approach to optimal parameter estimation (Springer).
24. Yang, J., Lee, S.H., Goddard, M.E., and Visscher, P.M. (2011). GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. *88*, 76–82.
25. Devlin, B., and Roeder, K. (1999). Genomic control for association studies. Biometrics *55*, 997–1004.
26. Lee, S.H., Harold, D., Nyholt, D.R., Goddard, M.E., Zondervan, K.T., Williams, J., Montgomery, G.W., Wray, N.R., and Visscher, P.M.; ANZGene Consortium; International Endogene Consortium; Genetic and Environmental Risk for Alzheimer's disease Consortium (2013). Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer's disease, multiple sclerosis and endometriosis. Hum. Mol. Genet. *22*, 832–841.
27. Lee, S.H., DeCandia, T.R., Ripke, S., Yang, J., Sullivan, P.F., Goddard, M.E., Keller, M.C., Visscher, P.M., and Wray, N.R.; Schizophrenia Psychiatric Genome-Wide Association Study Consortium (PGC-SCZ); International Schizophrenia Consortium (ISC); Molecular Genetics of Schizophrenia Collaboration (MGS) (2012). Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. Nat. Genet. *44*, 247–250.
28. Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C.C., Patsopoulos, N.A., Moutsianas, L., Dilthey, A., Su, Z., Freeman, C., Hunt, S.E., et al.; International Multiple Sclerosis Genetics Consortium; Wellcome Trust Case Control Consortium 2 (2011). Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature *476*, 214–219.
29. Golan, D., Lander, E.S., and Rosset, S. (2014). Measuring missing heritability: inferring the contribution of common variants. Proc. Natl. Acad. Sci. USA *111*, E5272–E5281.
30. Yang, J., Weedon, M.N., Purcell, S., Lettre, G., Estrada, K., Willer, C.J., Smith, A.V., Ingelsson, E., O'Connell, J.R., Mangino, M., et al.; GIANT Consortium (2011). Genomic inflation factors under polygenic inheritance. Eur. J. Hum. Genet. *19*, 807–812.
31. Bulik-Sullivan, B.K., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Patterson, N., Daly, M.J., Price, A.L., and Neale, B.M.; Schizophrenia Working Group of the Psychiatric Genomics

Consortium (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat. Genet. *47*, 291–295.
32. Visscher, P.M., Hill, W.G., and Wray, N.R. (2008). Heritability in the genomics era—concepts and misconceptions. Nat. Rev. Genet. *9*, 255–266.
33. Plomin, R., Haworth, C.M.A., and Davis, O.S.P. (2009). Common disorders are quantitative traits. Nat. Rev. Genet. *10*, 872–878.
34. Loh, P.-R., Tucker, G., Bulik-Sullivan, B.K., Vilhjálmsson, B.J., Finucane, H.K., Salem, R.M., Chasman, D.I., Ridker, P.M., Neale, B.M., Berger, B., et al. (2015). Efficient Bayesian mixed-model analysis increases association power in large cohorts. Nat. Genet. *47*, 284–290.
35. Zhou, X., and Stephens, M. (2014). Efficient multivariate linear mixed model algorithms for genome-wide association studies. Nat. Methods *11*, 407–409.

The American Journal of Human Genetics

Supplemental Data

Mixed Model with Correction for Case-Control Ascertainment Increases Association Power

Tristan J. Hayeck, Noah A. Zaitlen, Po-Ru Loh, Bjarni Vilhjalmsson, Samuela Pollack, Alexander Gusev, Jian Yang, Guo-Bo Chen, Michael E. Goddard, Peter M. Visscher, Nick Patterson, and Alkes L. Price

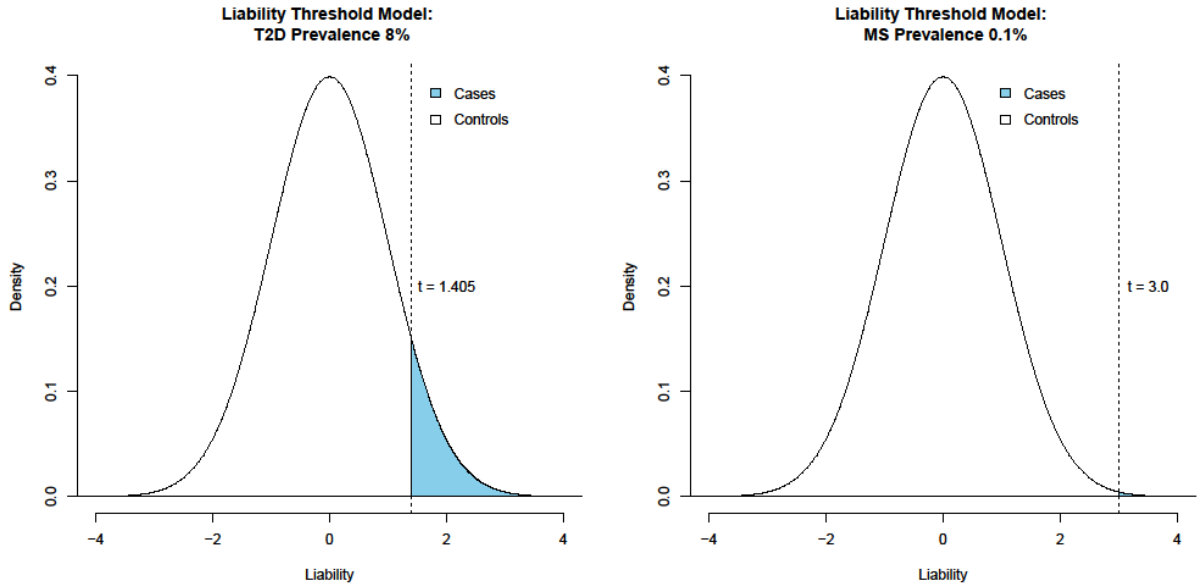

Figure S1. Liability Threshold Model. The liability threshold model performs a transformation based on disease prevalence. As ascertainment becomes more drastic so does the difference between the PML for cases versus controls. In Figure S1, the portion of the population above the threshold is a case (blue). For T2D, at a prevalence of 8% (blue), the threshold is set to 1.405. In this region, the expected value for the posterior liability is 1.85 and the expected value for the controls is -0.14. Comparing T2D to MS with disease prevalence around 0.1% and t around 3.00, the PML_{indiv} for a control is 0.00 and 3.33 for a case. As the disease prevalence goes down the difference in the PML_{indiv} for cases versus controls increases, the transformation plays a larger role for rare diseases and results in a power gain for the LTMLM.

A **B**

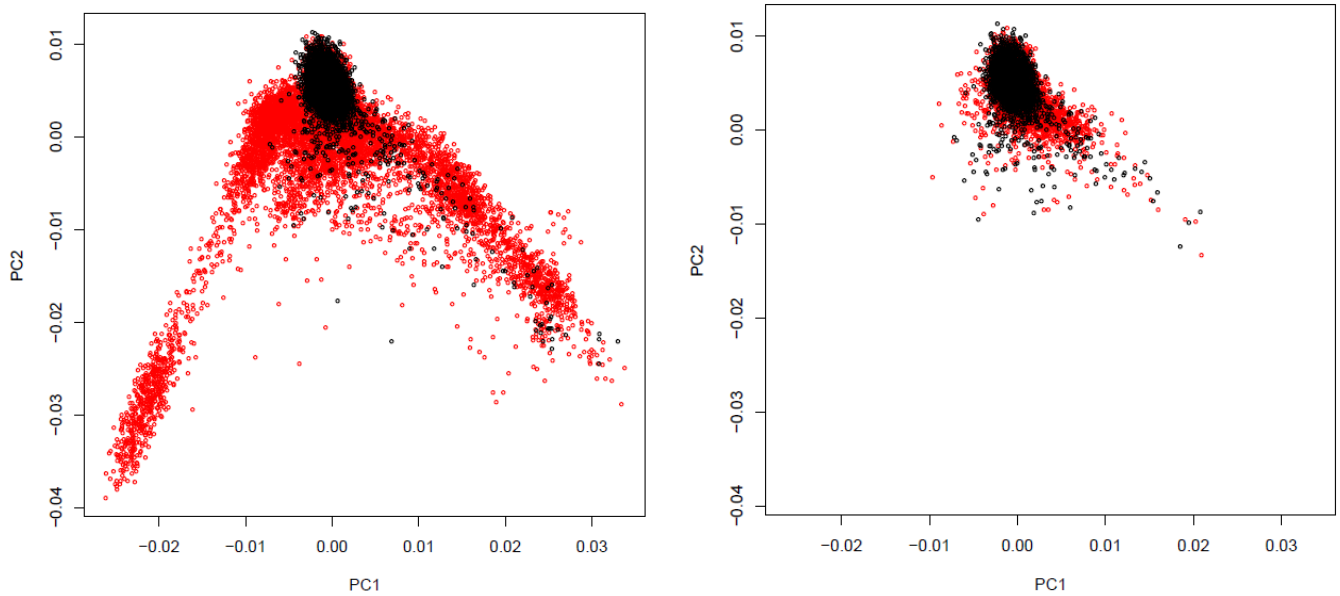

Figure S2. Mismatch in ancestry between MS cases and controls. We plot the first two principal components for (a) unmatched data with a severe mismatch (5,429 MS cases and 10,204controls), (b) stringently matched data using the first 20 PC(4,094 MS cases and 4,094 controls). The controls are depicted in red and cases in black. After PC matching the remaining samples show considerably less population stratification differentiation between cases and controls.

Term	Description
φ	Quantitative liability, the unobserved trait
β	Effect Size of the SNP
x	Genotype values of candidate SNP, normalized to mean 0 variance 1
u	Genetic random effect excluding the candidate SNP
e	Environmental component
X	Matrix of genotype values of non-candidate SNPs, normalized to mean 0 and variance 1
π	Observed binary case control phenotype.
t	Threshold corresponding to the disease prevalence
K	Prevalence of the disease in the population
P	Proportion of cases in the sample
$\widehat{\boldsymbol{\Theta}}$	Genetic Relationship Matrix (GRM) computed from the data
Θ	True underlying Genetic Relationship Matrix (GRM)
V	Phenotypic covariance matrix
I	Identity matrix
h^2	Heritability parameter

Table S1: Description of notation used and a brief description of the terms.

Computation	ATT	MLM	LTMLM
GRM and V^{-1}	NA	$O(MN^2)$	$O(MN^2)$
PML	NA	NA	$O(MN^2)$
Assoc. Statistic	$O(MN)$	$O(MN)$ or $O(MN^2)$	$O(MN)$
Overall	$O(MN)$	$O(MN^2)$	$O(MN^2)$

Table S2. Computational cost. M is the number of SNPs and N is the number of individuals. We assume that $M > N >$ #MCMC iterations. The details of the computational costs of MLM are provided in Table 1 of ref[11].

N	M	Prev	α	set	ATT Liability	MLM Liability	ATT	LogR	MLM	LTMLM
1000	1000	50%	0.05	causal	0.732(0.015)	0.741(0.015)	0.663(0.016)	0.663(0.016)	0.671(0.016)	0.670(0.016)
			0.001	causal	0.548(0.017)	0.562(0.017)	0.444(0.017)	0.443(0.017)	0.447(0.017)	0.444(0.017)
			$1x10^{-6}$	causal	0.331(0.016)	0.353(0.016)	0.227(0.014)	0.224(0.014)	0.236(0.014)	0.234(0.014)
			$5x10^{-8}$	causal	0.275(0.015)	0.287(0.015)	0.179(0.013)	0.175(0.013)	0.177(0.013)	0.178(0.013)
			0.05	null	0.050(0.001)	0.049(0.001)	0.051(0.001)	0.051(0.001)	0.051(0.001)	0.051(0.001)
			0.001	null	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(2x10^{-4})$	$0.001(2x10^{-4})$	$0.001(2x10^{-4})$	$0.001(2x10^{-4})$
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		25%	0.05	causal	0.704(0.015)	0.712(0.015)	0.648(0.016)	0.648(0.016)	0.652(0.016)	0.653(0.016)
			0.001	causal	0.544(0.017)	0.557(0.017)	0.456(0.017)	0.455(0.017)	0.460(0.017)	0.466(0.017)
			$1x10^{-6}$	causal	0.379(0.016)	0.383(0.016)	0.271(0.015)	0.268(0.015)	0.267(0.015)	0.272(0.015)
			$5x10^{-8}$	causal	0.318(0.016)	0.334(0.016)	0.219(0.014)	0.215(0.014)	0.226(0.014)	0.224(0.014)
			0.05	null	0.051(0.001)	0.051(0.001)	0.050(0.001)	0.049(0.001)	0.050(0.001)	0.049(0.001)
			0.001	null	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		10%	0.05	causal	0.761(0.014)	0.752(0.014)	0.708(0.015)	0.708(0.015)	0.708(0.015)	0.711(0.015)
			0.001	causal	0.589(0.016)	0.601(0.016)	0.51(0.017)	0.509(0.017)	0.514(0.017)	0.516(0.017)
			$1x10^{-6}$	causal	0.410(0.016)	0.417(0.016)	0.325(0.016)	0.323(0.016)	0.325(0.016)	0.330(0.016)
			$5x10^{-8}$	causal	0.362(0.016)	0.377(0.016)	0.266(0.015)	0.259(0.015)	0.268(0.015)	0.277(0.015)
			0.05	null	0.050(0.001)	0.051(0.001)	0.049(0.001)	0.049(0.001)	0.050(0.001)	0.050(0.001)
			0.001	null	$0.001(2x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		1%	0.05	causal	0.803(0.013)	0.796(0.013)	0.767(0.014)	0.767(0.014)	0.753(0.014)	0.769(0.014)
			0.001	causal	0.666(0.016)	0.665(0.016)	0.632(0.016)	0.632(0.016)	0.619(0.016)	0.638(0.016)
			$1x10^{-6}$	causal	0.548(0.016)	0.543(0.016)	0.496(0.017)	0.491(0.017)	0.477(0.017)	0.501(0.017)
			$5x10^{-8}$	causal	0.504(0.016)	0.500(0.016)	0.449(0.016)	0.447(0.016)	0.433(0.016)	0.451(0.016)
			0.05	null	0.050(0.001)	0.048(0.001)	0.050(0.001)	0.050(0.001)	0.049(0.001)	0.050(0.001)
			0.001	null	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)

		0.1%	0.05	causal	0.820(0.013)	0.811(0.013)	0.815(0.013)	0.815(0.013)	0.801(0.013)	0.818(0.013)
			0.001	causal	0.727(0.015)	0.689(0.015)	0.697(0.015)	0.697(0.015)	0.656(0.016)	0.695(0.015)
			$1x10^{-6}$	causal	0.585(0.016)	0.559(0.016)	0.551(0.016)	0.547(0.016)	0.510(0.017)	0.559(0.016)
			$5x10^{-8}$	causal	0.548(0.016)	0.511(0.016)	0.512(0.016)	0.510(0.016)	0.475(0.017)	**0.525(0.016)**
			0.05	null	0.050(0.001)	0.046(0.001)	0.050(0.001)	0.049(0.001)	0.048(0.001)	0.050(0.001)
			0.001	null	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
5000	1000	50%	0.05	causal	0.700(0.016)	0.700 (0.016)	0.642(0.017)	0.639(0.017)	0.637(0.017)	0.636(0.017)
			0.001	causal	0.526(0.017)	0.529(0.017)	0.422(0.017)	0.420(0.017)	0.424(0.017)	0.424(0.017)
			$1x10^{-6}$	causal	0.337(0.016)	0.341(0.016)	0.234(0.015)	0.231(0.015)	0.240(0.015)	0.241(0.015)
			$5x10^{-8}$	causal	0.290(0.015)	0.288(0.015)	0.19(0.014)	0.185(0.013)	0.191(0.014)	0.190(0.013)
			0.05	null	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)
			0.001	null	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		25%	0.05	causal	0.727(0.015)	0.723(0.015)	0.649(0.016)	0.649(0.016)	0.647(0.016)	0.646(0.016)
			0.001	causal	0.534(0.017)	0.537(0.017)	0.446(0.017)	0.444(0.017)	0.443(0.017)	0.445(0.017)
			$1x10^{-6}$	causal	0.369(0.016)	0.376(0.016)	0.271(0.015)	0.269(0.015)	0.270(0.015)	0.270(0.015)
			$5x10^{-8}$	causal	0.324(0.016)	0.324(0.016)	0.228(0.014)	0.227(0.014)	0.228(0.014)	0.228(0.014)
			0.05	null	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)
			0.001	null	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		10%	0.05	causal	0.741(0.015)	0.747(0.015)	0.691(0.016)	0.690(0.016)	0.694(0.016)	0.694(0.016)
			0.001	causal	0.579(0.016)	0.586(0.016)	0.518(0.017)	0.516(0.017)	0.520(0.017)	0.516(0.017)
			$1x10^{-6}$	causal	0.405(0.016)	0.403(0.016)	0.305(0.016)	0.303(0.016)	0.305(0.016)	0.307(0.016)
			$5x10^{-8}$	causal	0.349(0.016)	0.355(0.016)	0.248(0.015)	0.245(0.015)	0.249(0.015)	0.249(0.015)
			0.05	null	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)	0.051($4x10^{-4}$)	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)	0.051($4x10^{-4}$)

			0.001	null	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($7x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($7x10^{-5}$)
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		1%	0.05	causal	0.827(0.012)	0.834(0.012)	0.806(0.013)	0.806(0.013)	0.801(0.013)	0.807(0.013)
			0.001	causal	0.700(0.015)	0.700(0.015)	0.668(0.016)	0.664(0.016)	0.66(0.016)	0.663(0.016)
			$1x10^{-6}$	causal	0.555(0.016)	0.551(0.016)	0.499(0.017)	0.496(0.017)	0.494(0.017)	0.503(0.017)
			$5x10^{-8}$	causal	0.514(0.016)	0.509(0.016)	0.443(0.016)	0.441(0.016)	0.44(0.016)	0.446(0.016)
			0.05	null	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)	0.050($4x10^{-4}$)
			0.001	null	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		0.1%	0.05	causal	0.845(0.012)	0.854(0.012)	0.852(0.012)	0.852(0.012)	0.848(0.012)	0.852(0.012)
			0.001	causal	0.771(0.014)	0.770(0.014)	0.775(0.014)	0.775(0.014)	0.770(0.014)	0.770(0.014)
			$1x10^{-6}$	causal	0.689(0.015)	0.667(0.015)	0.657(0.016)	0.657(0.016)	0.631(0.016)	**0.688(0.015)**
			$5x10^{-8}$	causal	0.666(0.016)	0.654(0.016)	0.620(0.016)	0.593(0.016)	0.555(0.016)	0.590(0.016)
			0.05	null	0.049($4x10^{-4}$)	0.049($4x10^{-4}$)	0.051($4x10^{-4}$)	0.051($4x10^{-4}$)	0.049($4x10^{-4}$)	0.051($4x10^{-4}$)
			0.001	null	0.001($7x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
5000	5000	50%	0.05	causal	0.696(0.007)	0.703(0.007)	0.630(0.007)	0.629(0.007)	0.633(0.007)	0.632(0.007)
			0.001	causal	0.519(0.008)	0.530(0.008)	0.415(0.008)	0.415(0.008)	0.424(0.008)	0.424(0.008)
			$1x10^{-6}$	causal	0.332(0.007)	0.344(0.007)	0.229(0.006)	0.229(0.006)	0.236(0.007)	0.235(0.007)
			$5x10^{-8}$	causal	0.285(0.007)	0.295(0.007)	0.186(0.006)	0.184(0.006)	0.191(0.006)	0.190(0.006)
			0.05	null	0.0491($4x10^{-4}$)	0.049($4x10^{-4}$)	0.049($4x10^{-4}$)	0.049($4x10^{-4}$)	0.049($4x10^{-4}$)	0.049($4x10^{-4}$)
			0.001	null	0.001 ($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)	0.001($6x10^{-5}$)
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		25%	0.05	causal	0.694(0.007)	0.700(0.007)	0.628(0.007)	0.628(0.007)	0.632(0.007)	0.632(0.007)
			0.001	causal	0.512(0.008)	0.523(0.008)	0.427(0.008)	0.427(0.008)	0.434(0.008)	0.434(0.008)

		$1x10^{-6}$	causal	0.344(0.007)	0.355(0.007)	0.260(0.007)	0.260(0.007)	0.265(0.007)	0.265(0.007)
		$5x10^{-8}$	causal	0.297(0.007)	0.307(0.007)	0.211(0.006)	0.211(0.006)	0.215(0.006)	0.215(0.006)
		0.05	null	$0.050(4x10^{-4})$	$0.050(4x10^{-4})$	$0.050(4x10^{-4})$	$0.050(4x10^{-4})$	$0.050(4x10^{-4})$	$0.050(4x10^{-4})$
		0.001	null	$0.001\ (7x10^{-5})$	$0.001(6x10^{-5})$	$0.001\ (6x10^{-5})$	$0.001(6x10^{-5})$	$0.001(6x10^{-5})$	$0.001(6x10^{-5})$
		$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
	10%	0.05	causal	0.751(0.006)	0.758(0.006)	0.704(0.007)	0.703(0.007)	0.706(0.007)	0.709(0.007)
		0.001	causal	0.592(0.007)	0.597(0.007)	0.516(0.008)	0.516(0.008)	0.517(0.008)	0.520(0.008)
		$1x10^{-6}$	causal	0.428(0.007)	0.434(0.007)	0.335(0.007)	0.334(0.007)	0.335(0.007)	0.339(0.007)
		$5x10^{-8}$	causal	0.373(0.007)	0.384(0.007)	0.278(0.007)	0.277(0.007)	0.280(0.007)	0.285(0.007)
		0.05	null	$0.049(4x10^{-4})$	$0.049(4x10^{-4})$	$0.049(4x10^{-4})$	$0.049(4x10^{-4})$	$0.049(4x10^{-4})$	$0.050(4x10^{-4})$
		0.001	null	$0.001\ (6x10^{-5})$	$0.001(6x10^{-5})$	$0.001\ (6x10^{-5})$	$0.001(6x10^{-5})$	$0.001(6x10^{-5})$	$0.001(6x10^{-5})$
		$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
	1%	0.05	causal	0.805(0.006)	0.803(0.006)	0.775(0.006)	0.775(0.006)	0.773(0.006)	0.780(0.006)
		0.001	causal	0.668(0.007)	0.664(0.007)	0.626(0.007)	0.626(0.007)	0.618(0.007)	0.631(0.007)
		$1x10^{-6}$	causal	0.526(0.007)	0.516(0.007)	0.479(0.007)	0.479(0.007)	0.468(0.007)	0.485(0.007)
		$5x10^{-8}$	causal	0.482(0.007)	0.469(0.007)	0.429(0.007)	0.428(0.007)	0.415(0.007)	0.433(0.007)
		0.05	null	$0.050(4x10^{-4})$	$0.045(4x10^{-4})$	$0.050(4x10^{-4})$	$0.050(4x10^{-4})$	$0.050(4x10^{-4})$	$0.050(4x10^{-4})$
		0.001	null	$0.001(6x10^{-5})$	$0.001(5x10^{-5})$	$0.001(6x10^{-5})$	$0.001(6x10^{-5})$	$0.001(6x10^{-5})$	$0.001(6x10^{-5})$
		$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
	0.1%	0.05	causal	0.833(0.005)	0.812(0.006)	0.818(0.006)	0.818(0.006)	0.799(0.006)	0.821(0.006)
		0.001	causal	0.726(0.007)	0.697(0.007)	0.702(0.007)	0.702(0.007)	0.670(0.007)	0.705(0.007)
		$1x10^{-6}$	causal	0.599(0.007)	0.554(0.007)	0.562(0.007)	0.560(0.007)	0.516(0.007)	0.564(0.007)
		$5x10^{-8}$	causal	0.554(0.007)	0.510(0.007)	0.516(0.007)	0.516(0.007)	0.470(0.007)	0.519(0.007)
		0.05	null	$0.050(4x10^{-4})$	$0.0324(4x10^{-4})$	$0.050(4x10^{-4})$	$0.050(4x10^{-4})$	$0.040(4x10^{-4})$	$0.050(4x10^{-4})$
		0.001	null	$0.001\ (6x10^{-5})$	$3x10^{-4}(4x10^{-5})$	$0.001(6x10^{-5})$	$0.001(6x10^{-5})$	$0.001(5x10^{-5})$	$0.001(6x10^{-5})$

			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
5000	50000	50%	0.05	causal	0.702(0.007)	0.703(0.007)	0.631(0.007)	0.631(0.007)	0.631(0.007)	0.631(0.007)
			0.001	causal	0.518(0.008)	0.520(0.008)	0.418(0.008)	0.418(0.008)	0.419(0.008)	0.420(0.008)
			$1x10^{-6}$	causal	0.331(0.007)	0.333(0.007)	0.231(0.006)	0.230(0.006)	0.232(0.006)	0.231(0.006)
			$5x10^{-8}$	causal	0.278(0.007)	0.280(0.007)	0.183(0.006)	0.182(0.006)	0.183(0.006)	0.184(0.006)
			0.05	null	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$
			0.001	null	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		25%	0.05	causal	0.706(0.007)	0.710(0.007)	0.652(0.007)	0.652(0.007)	0.652(0.007)	0.652(0.007)
			0.001	causal	0.539(0.008)	0.540(0.008)	0.450(0.008)	0.449(0.008)	0.450(0.008)	0.450(0.008)
			$1x10^{-6}$	causal	0.362(0.007)	0.363(0.007)	0.259(0.007)	0.259(0.007)	0.262(0.007)	0.261(0.007)
			$5x10^{-8}$	causal	0.308(0.007)	0.312(0.007)	0.207(0.006)	0.206(0.006)	0.209(0.006)	0.209(0.006)
			0.05	null	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$
			0.001	null	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		10%	0.05	causal	0.751(0.006)	0.755(0.006)	0.726(0.007)	0.726(0.007)	0.725(0.007)	0.720(0.007)
			0.001	causal	0.629(0.007)	0.631(0.007)	0.548(0.008)	0.545(0.008)	0.555(0.007)	0.551(0.007)
			$1x10^{-6}$	causal	0.466(0.007)	0.474(0.007)	0.355(0.007)	0.355(0.007)	0.355(0.007)	0.355(0.007)
			$5x10^{-8}$	causal	0.381(0.007)	0.385(0.007)	0.272(0.007)	0.270(0.007)	0.274(0.007)	0.273(0.007)
			0.05	null	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$
			0.001	null	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		1%	0.05	causal	0.793(0.006)	0.795(0.006)	0.770(0.006)	0.770(0.006)	0.769(0.006)	0.770(0.006)
			0.001	causal	0.671(0.007)	0.671(0.007)	0.634(0.007)	0.634(0.007)	0.634(0.007)	0.637(0.007)
			$1x10^{-6}$	causal	0.526(0.007)	0.526(0.007)	0.477(0.007)	0.476(0.007)	0.476(0.007)	0.480(0.007)

		$5x10^{-8}$	causal	0.484(0.007)	0.485(0.007)	0.431(0.007)	0.431(0.007)	0.429(0.007)	0.432(0.007)
		0.05	null	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$
		0.001	null	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$
		$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
	0.1%	0.05	causal	0.835(0.005)	0.833(0.005)	0.821(0.006)	0.820(0.006)	0.817(0.006)	0.821(0.006)
		0.001	causal	0.719(0.007)	0.717(0.007)	0.697(0.007)	0.697(0.007)	0.688(0.007)	0.697(0.007)
		$1x10^{-6}$	causal	0.588(0.007)	0.587(0.007)	0.561(0.007)	0.560(0.007)	0.557(0.007)	0.560(0.007)
		$5x10^{-8}$	causal	0.547(0.007)	0.545(0.007)	0.517(0.007)	0.517(0.007)	0.508(0.007)	0.518(0.007)
		0.05	null	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$	$0.050(1x10^{-4})$
		0.001	null	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$
		$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)

Table S3. Percentage of SNPs achieving alpha levels for simulated genotypes and simulated phenotypes. We report the true positive and false positives at different α levels. For completeness, we also report ATT and MLM statistics computed using the underlying liability, where we again observe a loss in power for MLM at lower prevalence. In bold are the settings where LTMLM demonstrates at least a 5% power improvement over MLM.

N	M	Prev	Set	Statistic	ATT Liability	MLM Liability	ATT	LogR	MLM	LTMLM
1000	1000	50%	causal	average	25.716(1.091)	26.947(1.132)	16.847(0.715)	16.367(0.675)	17.254(0.735)	17.229(0.736)
			null	average	1.005(0.005)	1.003(0.005)	1.008(0.005)	1.006(0.005)	1.008(0.005)	1.007(0.005)
			all	average	1.252(0.014)	1.262(0.015)	1.166(0.010)	1.160(0.009)	1.171(0.010)	1.169(0.01)
			all	λ_{GC}	1.034(0.008)	1.033(0.008)	1.031(0.009)	1.031(0.009)	1.033(0.008)	1.032(0.008)
		25%	causal	average	28.716(1.229)	30.022(1.275)	19.112(0.824)	18.475(0.773)	19.553(0.849)	19.571(0.849)
			null	average	1.003(0.005)	1.004(0.005)	1.000(0.004)	0.998(0.004)	1.002(0.004)	1.001(0.004)
			all	average	1.280(0.016)	1.294(0.016)	1.181(0.011)	1.173(0.010)	1.187(0.011)	1.187(0.011)
			all	λ_{GC}	1.019(0.007)	1.021(0.007)	1.014(0.007)	1.013(0.007)	1.026(0.007)	1.02(0.007)
		10%	causal	average	34.541(1.501)	35.559(1.539)	23.777(1.021)	22.821(0.943)	24.084(1.053)	24.738(1.098)
			null	average	1.001(0.005)	1.004(0.005)	1.001(0.004)	0.999(0.004)	1.004(0.004)	1.003(0.004)
			all	average	1.336(0.019)	1.349(0.019)	1.229(0.013)	1.217(0.012)	1.235(0.014)	1.241(0.014)
			all	λ_{GC}	1.021(0.008)	1.028(0.008)	1.025(0.008)	1.025(0.008)	1.030(0.008)	1.029(0.008)
		1%	causal	average	55.732(2.12)	54.012(2.06)	42.316(1.565)	39.858(1.421)	40.611(1.539)	**45.336(1.756)**
			null	average	1.002(0.004)	0.988(0.004)	0.999(0.004)	0.997(0.004)	0.992(0.004)	0.999(0.004)
			all	average	1.550(0.028)	1.518(0.027)	1.412(0.021)	1.386(0.019)	1.388(0.020)	1.443(0.023)
			all	λ_{GC}	1.027(0.008)	1.012(0.007)	1.02(0.008)	1.019(0.008)	1.009(0.008)	1.018(0.008)
		0.1%	causal	average	73.159(2.847)	63.052(2.549)	58.946(2.157)	54.625(1.924)	51.507(2.045)	**65.991(2.656)**
			null	average	0.997(0.004)	0.974(0.004)	0.996(0.004)	0.994(0.004)	0.978(0.004)	0.996(0.004)
			all	average	1.718(0.037)	1.595(0.032)	1.575(0.029)	1.53(0.026)	1.483(0.026)	1.646(0.034)
			all	λ_{GC}	1.035(0.009)	1.008(0.007)	1.023(0.007)	1.023(0.007)	1.000(0.007)	1.014(0.008)
5000	1000	50%	causal	average	25.224(1.076)	25.503(1.084)	16.488(0.703)	16.029(0.666)	16.652(0.713)	16.625(0.714)
			null	average	1.002(0.002)	1.002(0.002)	1.000(0.002)	0.998(0.002)	1.000(0.002)	0.999(0.002)
			all	average	1.050(0.003)	1.051(0.003)	1.031(0.003)	1.028(0.003)	1.032(0.003)	1.031(0.003)
			all	λ_{GC}	1.006(0.003)	1.007(0.003)	1.005(0.003)	1.005(0.003)	1.006(0.003)	1.005(0.004)
		25%	causal	average	28.012(1.156)	28.279(1.163)	18.545(0.767)	17.985(0.726)	18.707(0.778)	18.757(0.783)
			null	average	1.004(0.002)	1.004(0.002)	1.002(0.002)	1.000(0.002)	1.002(0.002)	1.002(0.002)

			all	average	1.058(0.004)	1.059(0.004)	1.037(0.003)	1.034(0.003)	1.038(0.003)	1.037(0.003)
			all	λ_{GC}	1.010(0.003)	1.011(0.003)	1.003(0.003)	1.003(0.003)	1.004(0.003)	1.004(0.003)
		10%	causal	average	32.953(1.379)	33.282(1.395)	22.883(0.948)	22.028(0.886)	23.065(0.965)	23.497(0.996)
			null	average	1.005(0.002)	1.005(0.002)	1.004(0.002)	1.002(0.002)	1.004(0.002)	1.003(0.002)
			all	average	1.069(0.004)	1.070(0.004)	1.047(0.003)	1.044(0.003)	1.048(0.003)	1.048(0.003)
			all	λ_{GC}	1.009(0.003)	1.009(0.004)	0.999(0.003)	0.999(0.003)	1.003(0.003)	1.001(0.003)
		1%	causal	average	57.101(2.112)	56.546(2.100)	43.459(1.570)	40.939(1.426)	43.546(1.611)	**46.302(1.760)**
			null	average	1.002(0.002)	1.002(0.002)	1.001(0.002)	0.999(0.002)	1.001(0.002)	1.001(0.002)
			all	average	1.115(0.006)	1.113(0.006)	1.086(0.005)	1.079(0.004)	1.086(0.005)	1.091(0.005)
			all	λ_{GC}	1.007(0.004)	1.009(0.003)	1.003(0.003)	1.002(0.003)	1.005(0.003)	1.003(0.004)
		0.1%	causal	average	74.089(2.623)	68.943(2.51)	61.561(2.057)	57.182(1.828)	58.752(2.08)	**67.284(2.450)**
			null	average	0.997(0.002)	0.993(0.002)	0.996(0.002)	0.995(0.002)	0.991(0.002)	0.996(0.002)
			all	average	1.143(0.007)	1.129(0.007)	1.117(0.006)	1.107(0.005)	1.106(0.006)	1.128(0.007)
			all	λ_{GC}	1.004(0.001)	0.994(0.002)	0.989(0.001)	0.988(0.001)	1.005(0.002)	0.996(0.002)
5000	5000	50%	causal	average	25.372(0.496)	26.69(0.52)	16.492(0.325)	16.399(0.321)	16.88(0.332)	16.867(0.332)
			null	average	0.993(0.002)	0.994(0.002)	0.988(0.002)	0.988(0.002)	0.99(0.002)	0.989(0.002)
			all	average	1.237(0.006)	1.251(0.007)	1.143(0.004)	1.142(0.004)	1.148(0.004)	1.148(0.004)
			all	λ_{GC}	1.010(0.004)	1.01(0.003)	1.010(0.003)	1.010(0.003)	1.014(0.003)	1.014(0.003)
		25%	causal	average	28.119(0.588)	29.56(0.617)	18.637(0.388)	18.509(0.383)	19.014(0.396)	19.056(0.398)
			null	average	1.001(0.002)	1.001(0.002)	1.000(0.002)	1.000 (0.002)	1.000(0.002)	1.001(0.002)
			all	average	1.272(0.007)	1.287(0.008)	1.177(0.005)	1.175(0.005)	1.180(0.005)	1.181(0.005)
			all	λ_{GC}	1.019(0.004)	1.017(0.004)	1.012(0.004)	1.012(0.004)	1.013(0.004)	1.013(0.004)
		10%	causal	average	36.38(0.716)	37.626(0.742)	25.235(0.501)	25.014(0.492)	25.386(0.506)	25.778(0.514)
			null	average	0.996(0.002)	0.994(0.002)	0.993(0.002)	0.992(0.002)	0.991(0.002)	0.992(0.002)
			all	average	1.350(0.009)	1.361(0.009)	1.235(0.006)	1.233(0.006)	1.235(0.006)	1.240(0.007)
			all	λ_{GC}	1.016(0.003)	1.017(0.004)	1.005(0.003)	1.005(0.003)	1.007(0.003)	1.008(0.003)
		1%	causal	average	58.172(1.131)	55.473(1.077)	45.376(0.878)	44.682(0.852)	42.594(0.825)	**46.691(0.913)**
			null	average	1.000(0.002)	0.954(0.002)	1.000 (0.002)	0.999(0.002)	0.990(0.002)	1(0.002)

			all	average	1.571(0.014)	1.499(0.013)	1.444(0.011)	1.436(0.011)	1.406(0.010)	1.457(0.011)
			all	λ_{GC}	1.020(0.003)	0.970(0.005)	1.020(0.003)	1.020(0.003)	1.011(0.003)	1.019(0.003)
		0.1%	causal	average	81.745(1.565)	66.716(1.29)	68.648(1.301)	67.099(1.248)	56.303(1.082)	**70.81(1.364)**
			null	average	1.001(0.002)	0.840(0.002)	1.000(0.002)	1.000(0.002)	0.918(0.002)	1.000 (0.002)
			all	average	1.809(0.019)	1.499(0.016)	1.677(0.016)	1.661(0.016)	1.472(0.013)	1.698(0.017)
			all	λ_{GC}	1.022(0.003)	0.861(0.006)	1.026(0.003)	1.026(0.003)	0.942(0.005)	1.025(0.003)
5000	50000	50%	causal	average	25.501(0.504)	25.642(0.507)	16.624(0.331)	16.529(0.327)	16.673(0.332)	16.69(0.333)
			null	average	0.999(0.001)	0.999(0.001)	1.000 (0.001)	0.999(0.001)	1.000 (0.001)	1.000 (0.001)
			all	average	1.023(0.001)	1.024(0.001)	1.015(0.001)	1.015(0.001)	1.016(0.001)	1.015(0.001)
			all	λ_{GC}	0.999(0.001)	1.000 (0.001)	1.003(0.001)	1.002(0.001)	0.999(0.001)	0.999(0.001)
		25%	causal	average	28.475(0.555)	28.638(0.557)	18.965(0.37)	18.843(0.366)	19.03(0.372)	19.04(0.372)
			null	average	1.003(0.001)	1.003(0.001)	1.003(0.001)	1.003(0.001)	1.003(0.001)	1.003(0.001)
			all	average	1.030(0.001)	1.031(0.001)	1.021(0.001)	1.020(0.001)	1.021(0.001)	1.021(0.001)
			all	λ_{GC}	1.004(0.001)	1.004(0.001)	1.006(0.001)	1.006(0.001)	1.003(0.001)	1.003(0.001)
		10%	causal	average	34.358(0.639)	34.704(0.647)	23.71(0.444)	23.528(0.437)	23.868(0.449)	23.910(0.450)
			null	average	0.999(0.001)	1.000 (0.001)	1.001(0.001)	1.000 (0.001)	1.001(0.001)	1.001(0.001)
			all	average	1.033(0.001)	1.034(0.001)	1.023(0.001)	1.023(0.001)	1.024(0.001)	1.024(0.001)
			all	λ_{GC}	1.006(0.001)	1.008(0.002)	1.007(0.001)	1.007(0.001)	1.005(0.001)	1.004(0.001)
		1%	causal	average	60.112(1.139)	59.863(1.135)	46.683(0.883)	45.969(0.859)	46.44(0.881)	47.368(0.905)
			null	average	1.000 (0.001)	1.000 (0.001)	0.999(0.001)	0.999(0.001)	0.999(0.001)	0.999(0.001)
			all	average	1.059(0.002)	1.058(0.002)	1.045(0.001)	1.044(0.001)	1.045(0.001)	1.045(0.001)
			all	λ_{GC}	1.002(0.001)	1.002(0.001)	1.004(0.001)	1.004(0.001)	1.001(0.001)	1.000 (0.001)
		0.1%	causal	average	79.864(1.54)	77.754(1.502)	67.059(1.278)	65.561(1.225)	65.232(1.251)	**68.618(1.333)**
			null	average	1.000 (0.001)	0.999(0.001)	0.999(0.001)	0.999(0.001)	0.999(0.001)	0.999(0.001)
			all	average	1.078(0.002)	1.076(0.002)	1.065(0.002)	1.063(0.002)	1.063(0.002)	1.067(0.002)
			all	λ_{GC}	1.002(0.001)	1.001(0.001)	1.004(0.001)	1.004(0.001)	1.000(0.001)	1.000(0.001)

Table S4. Complete results on simulated genotypes and simulated phenotypes. Results are analogous to Table 2, but are reported for other values of *M* and *N* and consist of the same simulations as S3. For completeness, we also report ATT and MLM statistics computed using the underlying liability, where we again

observe a loss in power for MLM at lower prevalence. In bold are the settings where LTMLM demonstrates at least a 5% power improvement over MLM.

M	Prev	Set	Statistic	ATT Liability	MLM Liability	ATT	ATT+PCs	LogR	LogR +PCs	MLM	LTMLM
1000	10%	**causal**	average	36.835(1.515)	16.352(0.683)	24.736(1.002)	23.406(0.951)	23.622(0.93)	22.399(0.883)	13.222(0.553)	12.172(0.536)
		null	average	2.238(0.032)	0.789(0.014)	1.814(0.022)	1.632(0.020)	1.780(0.020)	1.601(0.019)	0.849(0.012)	0.753(0.011)
		all	average	2.584(0.036)	0.944(0.016)	2.043(0.025)	1.849(0.023)	1.999(0.023)	1.809(0.022)	0.973(0.013)	0.867(0.013)
		all	λ_{GC}	1.376(0.042)	0.373(0.006)	1.240(0.028)	1.064(0.010)	1.239(0.028)	1.062(0.010)	0.518(0.007)	0.458(0.011)
10000	10%	**causal**	average	33.100(1.375)	29.318(1.217)	22.606(0.943)	21.99(0.919)	21.672(0.876)	21.104(0.854)	20.412(0.867)	19.817(0.874)
		null	average	1.342(0.004)	0.965(0.003)	1.232(0.003)	1.084(0.003)	1.226(0.003)	1.078(0.002)	0.976(0.002)	0.948(0.002)
		all	average	1.374(0.004)	0.993(0.003)	1.253(0.003)	1.104(0.003)	1.246(0.003)	1.098(0.003)	0.995(0.003)	0.967(0.003)
		all	λ_{GC}	1.258(0.036)	0.882(0.003)	1.174(0.025)	1.023(0.003)	1.174(0.025)	1.023(0.003)	0.920(0.003)	0.893(0.013)
20000	10%	**causal**	average	34.93(1.481)	32.625(1.39)	24.098(1.037)	23.592(1.014)	23.031(0.963)	22.568(0.943)	22.892(1.003)	22.811(1.006)
		null	average	1.281(0.002)	0.982(0.002)	1.187(0.002)	1.042(0.001)	1.182(0.002)	1.038(0.001)	0.986(0.001)	0.981(0.001)
		all	average	1.298(0.002)	0.998(0.002)	1.198(0.002)	1.053(0.002)	1.193(0.002)	1.048(0.001)	0.997(0.002)	0.992(0.002)
		all	λ_{GC}	1.247(0.029)	0.939(0.002)	1.168(0.020)	1.013(0.002)	1.167(0.020)	1.013(0.002)	0.958(0.002)	0.953(0.006)

Table S5. Results on simulated genotypes and simulated phenotypes with population structure. We report average χ^2 statistics for simulations with population structure averaged across 100 simulations for each parameter setting (see main text).

N	M	Prev	α	set	ATT Liability	MLM Liability	ATT	ATT+PCs	LogR	LogR +PCs	MLM	LTMLM
1000	1000	10%	0.05	causal	0.758(0.014)	0.661(0.016)	0.719(0.015)	0.715(0.015)	0.719(0.015)	0.716(0.015)	0.611(0.017)	0.596(0.017)
			0.001	causal	0.623(0.016)	0.421(0.017)	0.541(0.017)	0.535(0.017)	0.540(0.017)	0.530(0.017)	0.36(0.017)	0.334(0.016)
			$1x10^{-6}$	causal	0.435(0.016)	0.220(0.014)	0.323(0.016)	0.316(0.016)	0.315(0.016)	0.307(0.016)	0.183(0.013)	0.153(0.013)
			$5x10^{-8}$	causal	0.373(0.016)	0.176(0.013)	0.274(0.015)	0.261(0.015)	0.265(0.015)	0.256(0.015)	0.143(0.012)	0.116(0.011)
			0.05	null	0.097(0.001)	0.020(0.001)	0.086(0.001)	0.069(0.001)	0.085(0.001)	0.069(0.001)	0.022(0.001)	0.020(0.001)
			0.001	null	0.022(0.001)	0.012(0.001)	0.018(0.001)	0.016(0.001)	0.018(0.001)	0.016(0.001)	0.011(0.001)	0.010(0.001)
			$1x10^{-6}$	null	$0.012(5x10^{-4})$	$0.006(4x10^{-4})$	0.010 ($4x10^{-4}$)	$0.009(4x10^{-4})$	0.009 ($4x10^{-4}$)	$0.009(4x10^{-4})$	$0.005(4x10^{-4})$	$0.004(3x10^{-4})$
			$5x10^{-8}$	null	$0.011(5x10^{-4})$	$0.005(4x10^{-4})$	0.008 ($4x10^{-4}$)	$0.008(4x10^{-4})$	$0.008(4x10^{-4})$	$0.007(4x10^{-4})$	$0.004(3x10^{-4})$	$0.003(3x10^{-4})$
1000	10000		0.05	causal	0.745(0.015)	0.729(0.015)	0.707(0.015)	0.702(0.016)	0.707(0.015)	0.702(0.016)	0.682(0.016)	0.674(0.016)
			0.001	causal	0.587(0.016)	0.557(0.017)	0.495(0.017)	0.492(0.017)	0.494(0.017)	0.492(0.017)	0.475(0.017)	0.451(0.017)
			$1x10^{-6}$	causal	0.411(0.016)	0.380(0.016)	0.322(0.016)	0.314(0.016)	0.316(0.016)	0.311(0.016)	0.295(0.016)	0.280(0.015)
			$5x10^{-8}$	causal	0.347(0.016)	0.315(0.016)	0.270(0.015)	0.269(0.015)	0.266(0.015)	0.259(0.015)	0.245(0.015)	0.225(0.014)
			0.05	null	$0.08(4x10^{-4})$	$0.038(3x10^{-4})$	0.071 ($4x10^{-4}$)	$0.054(3x10^{-4})$	$0.071(4x10^{-4})$	$0.053(3x10^{-4})$	$0.042(3x10^{-4})$	$0.04(3x10^{-4})$
			0.001	null	$0.006(1x10^{-4})$	$0.002(7x10^{-5})$	0.004 ($9x10^{-5}$)	$0.003(7x10^{-5})$	$0.004(9x10^{-5})$	$0.003(7x10^{-5})$	$0.002(6x10^{-5})$	$0.002(7x10^{-5})$
			$1x10^{-6}$	null	$0.001(5x10^{-5})$	$0.001(5x10^{-5})$	0.001 ($4x10^{-5}$)	$0.001(4x10^{-5})$	$0.001(4x10^{-5})$	$0.001(4x10^{-5})$	$0.001(4x10^{-5})$	$0.001(4x10^{-5})$
			$5x10^{-8}$	null	$0.001(5x10^{-5})$	$0.001(5x10^{-5})$	0.001 ($4x10^{-5}$)	$0.001(4x10^{-5})$	$0.001(4x10^{-5})$	$0.001(4x10^{-5})$	$0.001(4x10^{-5})$	$0.001(4x10^{-5})$
1000	20000		0.05	causal	0.762(0.014)	0.753(0.014)	0.690(0.016)	0.687(0.016)	0.690(0.016)	0.686(0.016)	0.686(0.016)	0.682(0.016)
			0.001	causal	0.588(0.016)	0.578(0.017)	0.512(0.017)	0.506(0.017)	0.511(0.017)	0.505(0.017)	0.490(0.017)	0.488(0.017)
			$1x10^{-6}$	causal	0.412(0.016)	0.400(0.016)	0.315(0.016)	0.309(0.016)	0.311(0.016)	0.306(0.016)	0.297(0.015)	0.300(0.016)
			$5x10^{-8}$	causal	0.356(0.016)	0.340(0.016)	0.264(0.015)	0.262(0.015)	0.260(0.015)	0.258(0.015)	0.254(0.015)	0.254(0.015)
			0.05	null	$0.078(3x10^{-4})$	$0.043(2x10^{-4})$	0.069 ($2x10^{-5}$)	$0.052(2x10^{-4})$	$0.068(2x10^{-4})$	$0.051(2x10^{-4})$	$0.046(2x10^{-4})$	$0.045(2x10^{-4})$
			0.001	null	$0.005(7x10^{-5})$	$0.002(4x10^{-5})$	0.003 ($6x10^{-5}$)	$0.002(4x10^{-5})$	$0.003(5x10^{-5})$	$0.002(4x10^{-5})$	$0.002(4x10^{-5})$	$0.002(4x10^{-5})$
			$1x10^{-6}$	null	$0.001(3x10^{-5})$	$0.001(3x10^{-5})$	0.001 ($2x10^{-5}$)	0 (NA)	0 (NA)	0 (NA)	0 (NA)	0 (NA)

	$5x10^{-8}$	null	$0.001(2x10^{-5})$	$0.001(2x10^{-5})$	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)

Table S6. Percentage of SNPs achieving alpha levels for simulated genotypes and simulated phenotypes with population structure. We report the true positive and false positives at different α levels for the same simulations as in Table S5 (see main text).

		Liability		Observed	
M	Prevalence	H-E	REML	H-E	REML
1000	50%	0.259(0.013)	0.252(0.01)	0.165(0.008)	0.161(0.006)
	25%	0.241(0.010)	0.238(0.008)	0.173(0.007)	0.171(0.006)
	10%	0.245(0.011)	0.242(0.007)	0.233(0.010)	0.230(0.007)
10000	50%	0.236(0.014)	0.245(0.013)	0.150(0.009)	0.156(0.008)
	25%	0.250(0.014)	0.264(0.013)	0.180(0.010)	0.190(0.010)
	10%	0.259(0.012)	0.261(0.009)	0.246(0.011)	0.248(0.009)

Table S7. Heritability parameter estimates on simulated genotypes and phenotypes. Results are analogous to Table 3, under different settings of *M* and *N*.

		Liability		Observed	
M	Prevalence	H-E	REML	H-E	REML
1000	10%	0.470(0.013)	0.415(0.005)	0.446(0.012)	0.394(0.005)
10000	10%	0.407(0.031)	0.526(0.011)	0.438(0.044)	0.500(0.010)
20000	10%	0.383(0.030)	0.531(0.013)	0.408(0.040)	0.505(0.012)

Table S8. Heritability parameter estimates on simulated genotypes and phenotypes with population structure. These results are from the same simulations used to generate Table S5 and S6. We report results on both liability and observed scales. The true h^2 explained by the SNPs used to build the GRM is 25% on the liability scale for all simulations.

N	M	True Prev	Specified Prev	Set	Statistic	ATT Liablity	MLM Liability	ATT	LogR	MLM	LTMLM
1000	1000	1%	1%	causal	average	55.732(2.12)	54.012(2.06)	42.316(1.565)	39.858(1.421)	40.611(1.539)	45.336(1.756)
				null	average	1.002(0.004)	0.988(0.004)	0.999(0.004)	0.997(0.004)	0.992(0.004)	0.999(0.004)
				all	average	1.550(0.028)	1.518(0.027)	1.412(0.021)	1.386(0.019)	1.388(0.020)	1.443(0.023)
				all	λGC	1.027(0.008)	1.012(0.007)	1.020(0.008)	1.019(0.008)	1.009(0.008)	1.018(0.008)
			3.4%	causal	average	As Above	As Above	As Above	As Above	As Above	44.441(1.805)
				null	average	As Above	As Above	As Above	As Above	As Above	1.009(0.005)
				all	average	As Above	As Above	As Above	As Above	As Above	1.444(0.023)
				all	λGC	As Above	As Above	As Above	As Above	As Above	1.035(0.008)
			0.2%	causal	average	As Above	As Above	As Above	As Above	As Above	45.532(1.851)
				null	average	As Above	As Above	As Above	As Above	As Above	1.002(0.004)
				all	average	As Above	As Above	As Above	As Above	As Above	1.448(0.024)
				all	λGC	As Above	As Above	As Above	As Above	As Above	1.023(0.008)
		0.1%	0.1%	causal	average	73.159(2.847)	63.052(2.549)	58.946(2.157)	54.625(1.924)	51.507(2.045)	65.991(2.656)
				null	average	0.997(0.004)	0.974(0.004)	0.996(0.004)	0.994(0.004)	0.978(0.004)	0.996(0.004)
				all	average	1.718(0.037)	1.595(0.032)	1.575(0.029)	1.530(0.026)	1.483(0.026)	1.646(0.034)
				all	λGC	1.035(0.009)	1.008(0.007)	1.023(0.007)	1.023(0.007)	1.000(0.007)	1.014(0.008)
			0.5%	causal	average	As Above	As Above	As Above	As Above	As Above	65.281(2.617)
				null	average	As Above	As Above	As Above	As Above	As Above	0.998(0.004)
				all	average	As Above	As Above	As Above	As Above	As Above	1.641(0.033)
				all	λGC	As Above	As Above	As Above	As Above	As Above	1.022(0.007)
			0.02%	causal	average	As Above	As Above	As Above	As Above	As Above	68.616(2.729)
				null	average	As Above	As Above	As Above	As Above	As Above	1.003(0.005)
				all	average	As Above	As Above	As Above	As Above	As Above	1.679(0.035)
				all	λGC	As Above	As Above	As Above	As Above	As Above	1.022(0.008)

Table S9. Simulated genotypes and phenotypes with mis-specification of the liability threshold. LTMLM was run at prevalence of 1% and 0.1% under mis-specification of the threshold, t=true +/- 0.5.

N	M	True Prev	Specified Prev	α	set	ATT Liability	LogR Liability	ATT	LogR	MLM	LTMLM
5000	5000	1%	1%	0.05	causal	0.803(0.013)	0.796(0.013)	0.767(0.014)	0.767(0.014)	0.753(0.014)	0.769(0.014)
				0.001	causal	0.666(0.016)	0.665(0.016)	0.632(0.016)	0.632(0.016)	0.619(0.016)	0.638(0.016)
				1x10-6	causal	0.548(0.016)	0.543(0.016)	0.496(0.017)	0.491(0.017)	0.477(0.017)	0.501(0.017)
				5x10-8	causal	0.504(0.016)	0.500(0.016)	0.449(0.016)	0.447(0.016)	0.433(0.016)	0.451(0.016)
				0.05	null	0.050(0.001)	0.048(0.001)	0.050(0.001)	0.050(0.001)	0.049(0.001)	0.050(0.001)
				0.001	null	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)
				1x10-6	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
				5x10-8	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
			3.4%	0.05	causal	As Above	As Above	As Above	As Above	As Above	0.783(0.014)
				0.001	causal	As Above	As Above	As Above	As Above	As Above	0.647(0.016)
				1x10-6	causal	As Above	As Above	As Above	As Above	As Above	0.469(0.017)
				5x10-8	causal	As Above	As Above	As Above	As Above	As Above	0.428(0.016)
				0.05	null	As Above	As Above	As Above	As Above	As Above	0.051(0.001)
				0.001	null	As Above	As Above	As Above	As Above	As Above	0.001($1x10^{-4}$)
				1x10-6	null	As Above	As Above	As Above	As Above	As Above	0(NA)
				5x10-8	null	As Above	As Above	As Above	As Above	As Above	0(NA)
			0.2%	0.05	causal	As Above	As Above	As Above	As Above	As Above	0.774(0.014)
				0.001	causal	As Above	As Above	As Above	As Above	As Above	0.638(0.016)
				1x10-6	causal	As Above	As Above	As Above	As Above	As Above	0.491(0.017)
				5x10-8	causal	As Above	As Above	As Above	As Above	As Above	0.439(0.016)
				0.05	null	As Above	As Above	As Above	As Above	As Above	0.050(0.001)
				0.001	null	As Above	As Above	As Above	As Above	As Above	0.001($1x10^{-4}$)
				1x10-6	null	As Above	As Above	As Above	As Above	As Above	0(NA)
				5x10-8	null	As Above	As Above	As Above	As Above	As Above	0(NA)
		0.1%	0.1%	0.05	causal	0.820(0.013)	0.811(0.013)	0.815(0.013)	0.815(0.013)	0.801(0.013)	0.818(0.013)
				0.001	causal	0.727(0.015)	0.689(0.015)	0.697(0.015)	0.697(0.015)	0.656(0.016)	0.695(0.015)
				1x10-6	causal	0.585(0.016)	0.559(0.016)	0.551(0.016)	0.547(0.016)	0.510(0.017)	0.559(0.016)

		5x10-8	causal	0.548(0.016)	0.511(0.016)	0.512(0.016)	0.510(0.016)	0.475(0.017)	0.525(0.016)
		0.05	null	0.050(0.001)	0.046(0.001)	0.050(0.001)	0.049(0.001)	0.048(0.001)	0.050(0.001)
		0.001	null	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)
		1x10-6	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
		5x10-8	null	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
	0.5%	0.05	causal	As Above	As Above	As Above	As Above	As Above	0.823(0.013)
		0.001	causal	As Above	As Above	As Above	As Above	As Above	0.717(0.015)
		1x10-6	causal	As Above	As Above	As Above	As Above	As Above	0.569(0.016)
		5x10-8	causal	As Above	As Above	As Above	As Above	As Above	0.526(0.016)
		0.05	null	As Above	As Above	As Above	As Above	As Above	0.049(0.001)
		0.001	null	As Above	As Above	As Above	As Above	As Above	0.001($2x10^{-4}$)
		1x10-6	null	As Above	As Above	As Above	As Above	As Above	0(NA)
		5x10-8	null	As Above	As Above	As Above	As Above	As Above	0(NA)
	0.02%	0.05	causal	As Above	As Above	As Above	As Above	As Above	0.824(0.012)
		0.001	causal	As Above	As Above	As Above	As Above	As Above	0.716(0.015)
		1x10-6	causal	As Above	As Above	As Above	As Above	As Above	0.573(0.016)
		5x10-8	causal	As Above	As Above	As Above	As Above	As Above	0.523(0.016)
		0.05	null	As Above	As Above	As Above	As Above	As Above	0.050(0.001)
		0.001	null	As Above	As Above	As Above	As Above	As Above	0.001($1x10^{-4}$)
		1x10-6	null	As Above	As Above	As Above	As Above	As Above	0(NA)
		5x10-8	null	As Above	As Above	As Above	As Above	As Above	0(NA)

Table S10. Percentage of SNPs achieving alpha levels for simulated genotypes and simulated phenotypes with mis-specification of the liability threshold.
This corresponds to the same set of simulations as table S9.

	Prevalence		Liability		Observed	
M	True	Specified	H-E	REML	H-E	REML
1000	1%	1%	0.246(0.007)	0.206(0.005)	0.445(0.013)	0.373(0.008)
		3.4%	0.350(0.012)	0.290(0.007)	0.462(0.016)	0.383(0.010)
		0.2%	0.182(0.006)	0.152(0.004)	0.445(0.014)	0.373(0.009)
	0.1%	0.1%	0.247(0.008)	0.175(0.004)	0.701(0.023)	0.498(0.011)
		0.5%	0.324(0.009)	0.234(0.004)	0.689(0.019)	0.499(0.009)
		0.02%	0.184(0.006)	0.132(0.003)	0.677(0.022)	0.486(0.011)

Table S11. Heritability parameter estimates on simulated genotypes and phenotypes with misspecification of the liability threshold. This corresponds to the same set of simulations as Table S9 and S10.

N	M	Prev	Set	Statistic	ATT	LogR	MLM	LTMLM
1000	1000	50%	causal	average	18.085(0.248)	17.859(0.242)	18.582(0.252)	18.576(0.254)
			null	average	0.997(0.004)	0.996(0.004)	0.998(0.004)	0.997(0.004)
			all	average	1.168(0.007)	1.164(0.007)	1.174(0.008)	1.173(0.008)
			all	λ_{GC}	1.023(0.007)	1.022(0.007)	1.026(0.008)	1.02(0.007)
		25%	causal	average	18.531(0.246)	18.295(0.24)	18.987(0.251)	18.954(0.253)
			null	average	1.005(0.005)	1.003(0.004)	1.005(0.005)	1.004(0.005)
			all	average	1.180(0.008)	1.176(0.007)	1.185(0.008)	1.183(0.008)
			all	λ_{GC}	1.037(0.007)	1.036(0.007)	1.036(0.007)	1.035(0.007)
		10%	causal	average	20.926(0.268)	20.628(0.26)	21.259(0.272)	21.347(0.272)
			null	average	1.005(0.004)	1.003(0.004)	1.005(0.004)	1.004(0.004)
			all	average	1.204(0.008)	1.200(0.008)	1.207(0.008)	1.207(0.008)
			all	λ_{GC}	1.031(0.007)	1.031(0.007)	1.036(0.008)	1.037(0.008)
		1%	causal	average	28.421(0.326)	27.886(0.314)	27.697(0.323)	28.974(0.338)
			null	average	1.009(0.005)	1.007(0.005)	1.008(0.005)	1.008(0.005)
			all	average	1.283(0.010)	1.276(0.010)	1.275(0.010)	1.288(0.010)
			all	λ_{GC}	1.027(0.008)	1.027(0.008)	1.025(0.008)	1.028(0.008)
		0.1%	causal	average	35.927(0.358)	35.100(0.342)	33.172(0.347)	36.654(0.373)
			null	average	0.999(0.004)	0.997(0.004)	0.992(0.004)	0.998(0.004)
			all	average	1.348(0.012)	1.338(0.012)	1.314(0.012)	1.355(0.013)
			all	λ_{GC}	1.031(0.007)	1.031(0.007)	1.012(0.007)	1.022(0.007)

Table S12. Simulated genotypes and phenotypes generated from a logit distribution. We report average χ^2 statistics for simulations with phenotypes generated from a logit distribution averaged across 100 simulations for each parameter setting (see main text).

N	M	Prev	α	set	ATT	LogR	MLM	LTMLM
1000	1000	50%	0.05	causal	0.988(0.003)	0.988(0.003)	0.988(0.003)	0.988(0.003)
			0.001	causal	0.827(0.012)	0.821(0.012)	0.832(0.012)	0.832(0.012)
			$1x10^{-6}$	causal	0.210(0.013)	0.198(0.013)	0.227(0.013)	0.230(0.013)
			$5x10^{-8}$	causal	0.077(0.008)	0.072(0.008)	0.089(0.009)	0.088(0.009)
			0.05	null	0.049(0.001)	0.049(0.001)	0.050(0.001)	0.050(0.001)
			0.001	null	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)
		25%	0.05	causal	0.997(0.002)	0.997(0.002)	0.997(0.002)	0.995(0.002)
			0.001	causal	0.832(0.012)	0.830(0.012)	0.849(0.011)	0.844(0.011)
			$1x10^{-6}$	causal	0.225(0.013)	0.209(0.013)	0.243(0.014)	0.245(0.014)
			$5x10^{-8}$	causal	0.090(0.009)	0.081(0.009)	0.098(0.009)	0.093(0.009)
			0.05	null	0.050(0.001)	0.050(0.001)	0.050(0.001)	0.050(0.001)
			0.001	null	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)
		10%	0.05	causal	0.997(0.002)	0.997(0.002)	0.998(0.001)	0.998(0.001)
			0.001	causal	0.893(0.010)	0.892(0.010)	0.894(0.010)	0.894(0.010)
			$1x10^{-6}$	causal	0.344(0.015)	0.334(0.015)	0.357(0.015)	0.365(0.015)
			$5x10^{-8}$	causal	0.140(0.011)	0.129(0.011)	0.154(0.011)	0.160(0.012)
			0.05	null	0.051(0.001)	0.051(0.001)	0.051(0.001)	0.050(0.001)
			0.001	null	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$	$0.001(1x10^{-4})$
			$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)
			$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)
		1%	0.05	causal	1(NA)	1(NA)	1(NA)	1(NA)
			0.001	causal	0.978(0.005)	0.977(0.005)	0.973(0.005)	0.980(0.004)
			$1x10^{-6}$	causal	0.640(0.015)	0.623(0.015)	0.613(0.015)	0.658(0.015)
			$5x10^{-8}$	causal	0.397(0.015)	0.378(0.015)	0.376(0.015)	0.419(0.016)

		0.05	null	0.051(0.001)	0.050(0.001)	0.051(0.001)	0.051(0.001)
		0.001	null	0.001($2x10^{-4}$)	0.001($2x10^{-4}$)	0.001($2x10^{-4}$)	0.001($2x10^{-4}$)
		$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)
		$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)
	0.1%	0.05	causal	1(NA)	1(NA)	1(NA)	1(NA)
		0.001	causal	0.996(0.002)	0.996(0.002)	0.995(0.002)	0.997(0.002)
		$1x10^{-6}$	causal	0.868(0.011)	0.860(0.011)	0.797(0.013)	0.880(0.01)
		$5x10^{-8}$	causal	0.685(0.015)	0.664(0.015)	0.584(0.016)	0.702(0.014)
		0.05	null	0.050(0.001)	0.050(0.001)	0.049(0.001)	0.050(0.001)
		0.001	null	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)	0.001($1x10^{-4}$)
		$1x10^{-6}$	null	0(NA)	0(NA)	0(NA)	0(NA)
		$5x10^{-8}$	null	0(NA)	0(NA)	0(NA)	0(NA)

Table S13. Percentage of SNPs achieving alpha levels for simulated genotypes and phenotypes generated from a logistic distribution. Results are the same simulations as described in Table S12

M	Prev	Set	Statistic	ATT Liability	MM Liability	ATT	ATT+PCs	LogR	LogR+PCs	MLM	LTMLM
1000	50%	causal	average	25.684(1.137)	29.055(1.26)	16.234(0.723)	15.775(0.705)	15.566(0.674)	15.14(0.658)	17.480(0.780)	17.412(0.775)
		null	average	1.698(0.015)	1.792(0.017)	1.444(0.011)	1.425(0.010)	1.433(0.010)	1.415(0.010)	1.472(0.011)	1.471(0.011)
		all	average	1.938(0.021)	2.065(0.023)	1.592(0.014)	1.569(0.013)	1.574(0.013)	1.552(0.013)	1.632(0.015)	1.630(0.015)
		all	λ_{GC}	1.274(0.018)	1.278(0.019)	1.214(0.017)	1.224(0.017)	1.214(0.017)	1.223(0.017)	1.234(0.017)	1.226(0.016)
	25%	causal	average	30.169(1.211)	33.771(1.343)	19.277(0.771)	18.581(0.754)	18.47(0.721)	17.822(0.706)	20.493(0.82)	20.642(0.831)
		null	average	1.821(0.017)	1.917(0.019)	1.551(0.012)	1.507(0.011)	1.538(0.012)	1.495(0.011)	1.571(0.012)	1.577(0.013)
		all	average	2.104(0.023)	2.236(0.025)	1.728(0.015)	1.678(0.014)	1.707(0.015)	1.658(0.014)	1.761(0.016)	1.768(0.016)
		all	λ_{GC}	1.290(0.018)	1.288(0.019)	1.256(0.016)	1.245(0.017)	1.255(0.016)	1.245(0.017)	1.241(0.015)	1.243(0.015)
	10%	causal	average	34.865(1.45)	37.861(1.551)	22.838(0.961)	21.406(0.878)	21.618(0.881)	20.336(0.808)	23.865(1.022)	24.661(1.064)
		null	average	1.973(0.020)	2.039(0.021)	1.664(0.014)	1.583(0.012)	1.647(0.014)	1.568(0.012)	1.668(0.014)	1.695(0.015)
		all	average	2.302(0.027)	2.397(0.028)	1.876(0.018)	1.781(0.016)	1.846(0.017)	1.756(0.015)	1.890(0.019)	1.925(0.019)
		all	λ_{GC}	1.301(0.017)	1.302(0.018)	1.271(0.016)	1.285(0.015)	1.270(0.016)	1.284(0.015)	1.251(0.017)	1.270(0.016)
10000	50%	causal	average	26.781(1.104)	27.44(1.129)	16.898(0.702)	16.815(0.698)	16.229(0.659)	16.156(0.655)	17.279(0.725)	17.26(0.725)
		null	average	1.114(0.002)	1.114(0.002)	1.078(0.002)	1.083(0.002)	1.074(0.002)	1.078(0.002)	1.078(0.002)	1.076(0.002)
		all	average	1.14(0.003)	1.140(0.003)	1.094(0.002)	1.099(0.002)	1.089(0.002)	1.094(0.002)	1.095(0.002)	1.092(0.002)
		all	λ_{GC}	1.046(0.005)	1.042(0.005)	1.039(0.005)	1.047(0.005)	1.039(0.005)	1.046(0.005)	1.039(0.005)	1.035(0.005)
	25%	causal	average	27.208(1.154)	27.938(1.178)	17.573(0.726)	17.293(0.696)	16.856(0.679)	16.61(0.654)	17.976(0.749)	18.077(0.758)
		null	average	1.112(0.002)	1.11(0.002)	1.076(0.002)	1.078(0.002)	1.071(0.002)	1.073(0.002)	1.074(0.002)	1.073(0.002)
		all	average	1.138(0.003)	1.137(0.003)	1.092(0.002)	1.094(0.002)	1.087(0.002)	1.089(0.002)	1.091(0.002)	1.090(0.002)
		all	λ_{GC}	1.048(0.005)	1.043(0.005)	1.040(0.004)	1.045(0.004)	1.040(0.004)	1.044(0.004)	1.040(0.004)	1.039(0.004)
	10%	causal	average	35.911(1.500)	36.598(1.525)	24.379(1.026)	24.116(1.014)	23.127(0.944)	22.894(0.934)	24.987(1.071)	25.399(1.091)
		null	average	1.161(0.003)	1.156(0.003)	1.112(0.002)	1.115(0.002)	1.107(0.002)	1.110(0.002)	1.108(0.002)	1.111(0.002)
		all	average	1.196(0.003)	1.192(0.003)	1.136(0.003)	1.138(0.003)	1.129(0.002)	1.132(0.002)	1.131(0.003)	1.135(0.003)
		all	λ_{GC}	1.059(0.005)	1.050(0.004)	1.051(0.005)	1.059(0.004)	1.050(0.005)	1.058(0.004)	1.044(0.004)	1.047(0.004)

Table S14. Complete results on real genotypes and simulated phenotypes. Results include results from Table 4 but we also report ATT and MLM statistics computed using the underlying liability. We report average χ^2 statistics for simulations with real genotypes and simulated phenotypes averaged across 100 simulations for each parameter setting.

M	Prev	α	ATT Liability	MLM Liability	ATT	ATT+PCs	LogR	LogR +PCs	MLM	LTMLM
1000	50%	0.05	0.689(0.016)	0.717(0.015)	0.614(0.017)	0.602(0.017)	0.614(0.017)	0.603(0.017)	0.630(0.017)	0.625(0.017)
		0.001	0.504(0.017)	0.533(0.017)	0.409(0.017)	0.408(0.017)	0.405(0.017)	0.403(0.017)	0.415(0.017)	0.415(0.017)
		$1x10^{-6}$	0.320(0.016)	0.348(0.016)	0.235(0.015)	0.227(0.014)	0.229(0.014)	0.218(0.014)	0.256(0.015)	0.256(0.015)
		$5x10^{-8}$	0.289(0.015)	0.309(0.016)	0.185(0.013)	0.178(0.013)	0.179(0.013)	0.172(0.013)	0.205(0.014)	0.200(0.014)
		0.05	0.095(0.001)	0.099(0.001)	0.084(0.001)	0.083(0.001)	0.084(0.001)	0.083(0.001)	0.085(0.001)	0.085(0.001)
		0.001	0.002(0.001)	0.022(0.001)	0.014(0.001)	$0.013(5x10^{-4})$	0.013(0.001)	$0.013(5x10^{-4})$	0.014(0.001)	0.014(0.001)
		$1x10^{-6}$	$0.007(4x10^{-4})$	$0.009(4x10^{-4})$	$0.003(3x10^{-4})$	$0.003(2x10^{-4})$	$0.003(3x10^{-4})$	$0.003(2x10^{-4})$	$0.004(3x10^{-4})$	$0.004(3x10^{-4})$
		$5x10^{-8}$	$0.005(3x10^{-4})$	$0.006(3x10^{-4})$	$0.002(2x10^{-4})$	$0.002(2x10^{-4})$	$0.002(2x10^{-4})$	$0.002(2x10^{-4})$	$0.003(2x10^{-4})$	$0.003(2x10^{-4})$
	25%	0.05	0.737(0.015)	0.750(0.014)	0.682(0.016)	0.671(0.016)	0.680(0.016)	0.671(0.016)	0.689(0.016)	0.692(0.016)
		0.001	0.579(0.017)	0.603(0.016)	0.491(0.017)	0.482(0.017)	0.486(0.017)	0.479(0.017)	0.498(0.017)	0.496(0.017)
		$1x10^{-6}$	0.405(0.016)	0.437(0.017)	0.294(0.016)	0.278(0.015)	0.281(0.015)	0.266(0.015)	0.314(0.016)	0.311(0.016)
		$5x10^{-8}$	0.336(0.016)	0.387(0.016)	0.231(0.014)	0.216(0.014)	0.218(0.014)	0.208(0.014)	0.248(0.015)	0.251(0.015)
		0.05	0.102(0.001)	0.104(0.001)	0.091(0.001)	0.09(0.001)	0.091(0.001)	0.090(0.001)	0.091(0.001)	0.092(0.001)
		0.001	0.023(0.001)	0.026(0.001)	0.016(0.001)	0.015(0.001)	0.016(0.001)	0.015(0.001)	0.017(0.001)	0.018(0.001)
		$1x10^{-6}$	$0.009(4x10^{-4})$	$0.01(4x10^{-4})$	$0.005(3x10^{-4})$	$0.004(3x10^{-4})$	$0.005(3x10^{-4})$	$0.004(3x10^{-4})$	$0.005(3x10^{-4})$	$0.005(3x10^{-4})$
		$5x10^{-8}$	$0.006(3x10^{-4})$	$0.007(4x10^{-4})$	$0.003(3x10^{-4})$	$0.003(2x10^{-4})$	$0.003(2x10^{-4})$	$0.003(2x10^{-4})$	$0.004(3x10^{-4})$	$0.004(3x10^{-4})$
	10%	0.05	0.730(0.015)	0.740(0.015)	0.678(0.016)	0.678(0.016)	0.678(0.016)	0.678(0.016)	0.684(0.016)	0.694(0.016)
		0.001	0.575(0.017)	0.593(0.016)	0.498(0.017)	0.492(0.017)	0.494(0.017)	0.490(0.017)	0.510(0.017)	0.517(0.017)
		$1x10^{-6}$	0.410(0.016)	0.435(0.017)	0.332(0.016)	0.328(0.016)	0.324(0.016)	0.316(0.016)	0.340(0.016)	0.343(0.016)
		$5x10^{-8}$	0.374(0.016)	0.388(0.016)	0.271(0.015)	0.265(0.015)	0.261(0.015)	0.251(0.015)	0.276(0.015)	0.281(0.015)
		0.05	0.106(0.001)	0.107(0.001)	0.096(0.001)	0.095(0.001)	0.095(0.001)	0.094(0.001)	0.095(0.001)	0.097(0.001)
		0.001	0.027(0.001)	0.028(0.001)	0.019(0.001)	0.016(0.001)	0.019(0.001)	0.016(0.001)	0.020(0.001)	0.020(0.001)
		$1x10^{-6}$	$0.01(4x10^{-4})$	$0.011(5x10^{-4})$	$0.007(4x10^{-4})$	$0.005(3x10^{-4})$	$0.006(3x10^{-4})$	$0.005(3x10^{-4})$	$0.007(4x10^{-4})$	$0.007(4x10^{-4})$
		$5x10^{-8}$	$0.008(4x10^{-4})$	$0.008(4x10^{-4})$	$0.005(3x10^{-4})$	$0.004(3x10^{-4})$	$0.005(3x10^{-4})$	$0.003(3x10^{-4})$	$0.005(3x10^{-4})$	$0.005(3x10^{-4})$
10000	50%	0.05	0.713(0.015)	0.72(0.015)	0.653(0.016)	0.652(0.016)	0.652(0.016)	0.651(0.016)	0.656(0.016)	0.657(0.016)
		0.001	0.535(0.017)	0.551(0.017)	0.427(0.017)	0.428(0.017)	0.422(0.017)	0.426(0.017)	0.432(0.017)	0.433(0.017)
		$1x10^{-6}$	0.356(0.016)	0.365(0.016)	0.245(0.015)	0.244(0.015)	0.238(0.015)	0.237(0.015)	0.248(0.015)	0.245(0.015)
		$5x10^{-8}$	0.306(0.015)	0.311(0.016)	0.187(0.013)	0.191(0.013)	0.181(0.013)	0.178(0.013)	0.190(0.013)	0.189(0.013)
		0.05	$0.058(3x10^{-4})$	$0.058(3x10^{-4})$	$0.057(3x10^{-4})$	$0.057(3x10^{-4})$	$0.056(3x10^{-4})$	$0.057(3x10^{-4})$	$0.057(3x10^{-4})$	$0.056(3x10^{-4})$

		0.001	0.004($9x10^{-5}$)	0.004($9x10^{-5}$)	0.003($7x10^{-5}$)	0.003($7x10^{-5}$)	0.003($7x10^{-5}$)	0.003($7x10^{-5}$)	0.003($8x10^{-5}$)	0.003($8x10^{-5}$)
		$1x10^{-6}$	0.001($4x10^{-5}$)	0.001($4x10^{-5}$)	0.001($3x10^{-5}$)	0.001($3x10^{-5}$)	0.001($3x10^{-5}$)	0.001($3x10^{-5}$)	0.001($3x10^{-5}$)	0.001($3x10^{-5}$)
		$5x10^{-8}$	0.001($4x10^{-5}$)	0.001($4x10^{-5}$)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
	25%	0.05	0.725(0.015)	0.726(0.015)	0.635(0.017)	0.633(0.017)	0.635(0.017)	0.632(0.017)	0.635(0.017)	0.635(0.017)
		0.001	0.529(0.017)	0.538(0.017)	0.440(0.017)	0.438(0.017)	0.438(0.017)	0.437(0.017)	0.448(0.017)	0.445(0.017)
		$1x10^{-6}$	0.352(0.016)	0.362(0.016)	0.266(0.015)	0.261(0.015)	0.256(0.015)	0.254(0.015)	0.270(0.015)	0.267(0.015)
		$5x10^{-8}$	0.303(0.015)	0.309(0.016)	0.209(0.014)	0.206(0.014)	0.198(0.014)	0.197(0.014)	0.215(0.014)	0.216(0.014)
		0.05	0.057($3x10^{-4}$)	0.057($3x10^{-4}$)	0.056($3x10^{-4}$)	0.056($3x10^{-4}$)	0.055($3x10^{-4}$)	0.056($3x10^{-4}$)	0.056($3x10^{-4}$)	0.055($3x10^{-4}$)
		0.001	0.004($9x10^{-5}$)	0.004($9x10^{-5}$)	0.003($7x10^{-5}$)	0.003($7x10^{-5}$)	0.003($7x10^{-5}$)	0.003($7x10^{-5}$)	0.003($7x10^{-5}$)	0.003($7x10^{-5}$)
		$1x10^{-6}$	0.001($5x10^{-5}$)	0.001($5x10^{-5}$)	0.001($3x10^{-5}$)	0.001($3x10^{-5}$)	0.001($3x10^{-5}$)	0.001($3x10^{-5}$)	0.001($3x10^{-5}$)	0.001($3x10^{-5}$)
		$5x10^{-8}$	0.001($4x10^{-5}$)	0.001($4x10^{-5}$)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
	10%	0.05	0.752(0.014)	0.756(0.014)	0.711(0.015)	0.711(0.015)	0.709(0.015)	0.706(0.015)	0.713(0.015)	0.715(0.015)
		0.001	0.597(0.016)	0.603(0.016)	0.519(0.017)	0.528(0.017)	0.518(0.017)	0.523(0.017)	0.520(0.017)	0.517(0.017)
		$1x10^{-6}$	0.437(0.017)	0.438(0.017)	0.329(0.016)	0.330(0.016)	0.322(0.016)	0.322(0.016)	0.328(0.016)	0.336(0.016)
		$5x10^{-8}$	0.380(0.016)	0.388(0.016)	0.281(0.015)	0.275(0.015)	0.266(0.015)	0.264(0.015)	0.283(0.015)	0.284(0.015)
		0.05	0.061($3x10^{-4}$)	0.06($3x10^{-4}$)	0.059($3x10^{-4}$)	0.059($3x10^{-4}$)	0.058($3x10^{-4}$)	0.059($3x10^{-4}$)	0.058($3x10^{-4}$)	0.058($3x10^{-4}$)
		0.001	0.005($1x10^{-4}$)	0.005($1x10^{-4}$)	0.004($8x10^{-5}$)	0.004($8x10^{-5}$)	0.003($8x10^{-5}$)	0.003($8x10^{-5}$)	0.004($8x10^{-5}$)	0.004($8x10^{-5}$)
		$1x10^{-6}$	0.001($5x10^{-5}$)	0.001($5x10^{-5}$)	0.001($4x10^{-5}$)	0.001($4x10^{-5}$)	0.001($4x10^{-5}$)	0.001($4x10^{-5}$)	0.001($4x10^{-5}$)	0.001($4x10^{-5}$)
		$5x10^{-8}$	0.001($5x10^{-5}$)	0.001($5x10^{-5}$)	0.001($4x10^{-5}$)	0.001($3x10^{-5}$)	0.001($4x10^{-5}$)	0.001($3x10^{-5}$)	0.001($4x10^{-5}$)	0.001($4x10^{-5}$)

Table S15. Percentage of SNPs achieving alpha levels for real genotypes and simulated phenotypes. Results include results for Table 4, but we also report results for ATT and MLM computed using the underlying liability.

Category	Controlling λ_{GC}	α	ATT	ATT+PCs	LogR	LogR+PCs	MLM	LTMLM
Genome Wide	No	0.05	0.093($5x10^{-4}$)	0.068($4x10^{-4}$)	0.093($5x10^{-4}$)	0.067($4x10^{-4}$)	0.067($4x10^{-4}$)	0.07($4x10^{-4}$)
Genome Wide	No	0.001	0.006($1x10^{-4}$)	0.003($9x10^{-5}$)	0.006($1x10^{-4}$)	0.003($9x10^{-5}$)	0.003($9x10^{-5}$)	0.003($1x10^{-4}$)
Genome Wide	No	$1x10^{-6}$	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
Genome Wide	No	$5x10^{-8}$	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
Published	No	0.05	0.747(0.050)	0.733(0.051)	0.747(0.050)	0.747(0.050)	0.707(0.053)	0.747(0.05)
Published	No	0.001	0.427(0.057)	0.387(0.056)	0.427(0.057)	0.387(0.056)	0.387(0.056)	0.360(0.055)
Published	No	$1x10^{-6}$	0.107(0.036)	0.080(0.031)	0.107(0.036)	0.067(0.029)	0.080(0.031)	0.093(0.034)
Published	No	$5x10^{-8}$	0.053(0.026)	0.040(0.023)	0.053(0.026)	0.027(0.019)	0.027(0.019)	0.053(0.026)
Genome Wide	Yes	0.05	0.053($4x10^{-4}$)	0.053($4x10^{-4}$)	0.053($4x10^{-4}$)	0.053($4x10^{-4}$)	0.053($4x10^{-4}$)	0.053($4x10^{-4}$)
Genome Wide	Yes	0.001	0.002($7x10^{-5}$)	0.002($7x10^{-5}$)	0.002($7x10^{-5}$)	0.002($7x10^{-5}$)	0.002($7x10^{-5}$)	0.002($7x10^{-5}$)
Genome Wide	Yes	$1x10^{-6}$	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
Genome Wide	Yes	$5x10^{-8}$	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)	0(NA)
Published	Yes	0.05	0.720(0.052)	0.693(0.053)	0.720(0.052)	0.707(0.053)	0.693(0.053)	0.733(0.051)
Published	Yes	0.001	0.320(0.054)	0.347(0.055)	0.320(0.054)	0.333(0.054)	0.347(0.055)	0.347(0.055)
Published	Yes	$1x10^{-6}$	0.040(0.023)	0.053(0.026)	0.040(0.023)	0.067(0.029)	0.040(0.023)	0.053(0.026)
Published	Yes	$5x10^{-8}$	0.013(0.013)	0.013(0.013)	0.013(0.013)	0.013(0.013)	0.013(0.013)	0.013(0.013)

Table S16. Proportion of SNPs achieving alpha levels for WTCCC2 MS data set. The number of known associated SNPs that are significant for LTMLM but not MLM (or vice versa) after controlling for λ_{GC} are 3(0) at $\alpha = 0.05$, 1(1) at $\alpha = 0.001$, 1(0) at $\alpha = 1x10^{-6}$, 0(0) at $\alpha = 5x10^{-8}$.

SNP Set	ATT	PCA	MLM	LTMLM
Genome Wide AVERAGE	1.38	1.16	1.14	1.17
Genome Wide LD Score INTERCEPT	1.29	1.09	1.08	1.10
Published SNPs AVERAGE	11.64	9.97	9.92	10.59
Published SNPs/Genome Wide Average	8.44	8.61	8.67	9.03
Published SNPs/LD score INTERCEPT	9.06	9.17	9.20	9.66

Table S17: Results on WTCCC2 MS data set with calibration via LD Score regression. We report the genome wide χ^2 averages using 10,034 individuals over 360,557 SNPs and the average across 75 published SNPs standardized by the genome wide average and LD Score regression intercept.

SNP Set	N	ATT	MLM	LTMLM	LTMLM REML
Genome Wide AVERAGE	8188	1.16	1.11	1.14	1.14
Published SNPs AVERAGE		8.94	8.26	8.76	8.82
Published SNPs/Genome Wide Average		7.73	7.45	7.71	7.73
Genome Wide AVERAGE	10034	1.38	1.14	1.17	1.24
Published SNPs AVERAGE		11.64	9.92	10.59	11.04
Published SNPs/Genome Wide Average		8.44	8.67	9.03	8.92
Genome Wide AVERAGE	15633	3.95	1.23	1.08	1.50
Published SNPs AVERAGE		18.54	11.30	5.76	13.76
Published SNPs/Genome Wide Average		4.69	9.20	5.32	9.15

Table S18: Results on WTCCC2 MS data set at different levels of QC. We report results for stringently matched (N = 8,188), partially matched (N = 10,034) and unmatched (N = 15,633) data sets (see main text). The additional column is for the LTMLM REML statistic calculated using the REML estimate of h^2. LTMLM using the REML estimate for h^2 produces inflated test statistics and it is not recommended.

	Liability		Observed	
N	HE	REML	HE	REML
8188	0.363 (0.0017)	0.260 (0.001)	0.979 (0.005)	0.702 (0.003)
10034	0.704 (0.009)	0.279 (0.001)	1.901 (0.025)	0.753 (0.002)
15633	2.792 (0.010)	0.293 (0.001)	7.543 (0.0266)	0.792 (0.002)

Table S19: Heritability parameter estimates on WTCCC2 MS data set at different levels of QC. We report results for stringently matched (N = 8,188), partially matched (N = 10,034) and unmatched (N = 15,633) data sets (see main text).