An Analytic Solution to the Computation of Power and Sample Size for Genetic Association Studies under a Pleiotropic Mode of Inheritance

Derek Gordon^{a, b} Douglas Londono^{a, b} Payal Patel^a Wonkuk Kim^d
Stephen J. Finch^c Gary A. Heiman^{a, b}

^aDepartment of Genetics and ^bHuman Genetics Institute, Rutgers, The State University of New Jersey, Piscataway, NJ, and ^cDepartment of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY, USA; ^dDepartment of Applied Statistics, Chung-Ang University, Seoul, Republic of Korea

Keywords

Pleiotropy · Multiple phenotypes · Genome-wide association study · Noncentrality parameter · Statistics · Method

Abstract

Our motivation here is to calculate the power of 3 statistical tests used when there are genetic traits that operate under a pleiotropic mode of inheritance and when qualitative phenotypes are defined by use of thresholds for the multiple quantitative phenotypes. Specifically, we formulate a multivariate function that provides the probability that an individual has a vector of specific quantitative trait values conditional on having a risk locus genotype, and we apply thresholds to define qualitative phenotypes (affected, unaffected) and compute penetrances and conditional genotype frequencies based on the multivariate function. We extend the analytic power and minimum-sample-size-necessary (MSSN) formulas for 2 categorical data-based tests (genotype, linear trend test [LTT]) of genetic association to the pleiotropic model. We further compare the MSSN of the genotype test and the LTT with that of a multivariate ANOVA (Pillai). We approximate the MSSN for statistics by linear models using a factorial design and ANOVA. With ANOVA decomposition,

we determine which factors most significantly change the power/MSSN for all statistics. Finally, we determine which test statistics have the smallest MSSN. In this work, MSSN calculations are for 2 traits (bivariate distributions) only (for illustrative purposes). We note that the calculations may be extended to address any number of traits. Our key findings are that the genotype test usually has lower MSSN requirements than the LTT. More inclusive thresholds (top/bottom 25% vs. top/bottom 10%) have higher sample size requirements. The Pillai test has a much larger MSSN than both the genotype test and the LTT, as a result of sample selection. With these formulas, researchers can specify how many subjects they must collect to localize genes for pleiotropic phenotypes.

© 2017 S. Karger AG, Basel

Introduction

In his review article on 100 years of pleiotropy, Stearns credits the Swiss geneticist Ludwig Plate as being the first to use the term in 1910 [1]. Stearns' definition was, "Plei-

D.G. and D.L. are co-first authors and contributed equally to this paper.

otropy refers to the phenomenon in which a single locus affects two or more apparently unrelated phenotypic traits and is often identified as a single mutation that affects two or more wild-type traits.” [1] We translate this definition into a mathematical model in the Methods section.

As of this writing, searching the term “pleiotropy” under “Topic” in the ISI Web of Science database yields over 11,000 publications. This number suggests that pleiotropy is both a common phenomenon and one that has been well studied. A significant number of these publications (over 1,300 according to ISI Web of Science) deal with mice, flies, plants, dogs, chickens, and other animals/organisms. There are a host of statistically powerful techniques available for gene mapping in these model organisms [see, e.g., 2 for mice].

In humans, there are numerous examples of pleiotropic effects that are correlated with traits and/or diseases. Some examples include colorectal cancer [3, 4], Crohn disease [3, 5–10], Alzheimer disease [11–19], and Marfan syndrome [20–22]. Papers by Baumgartner et al. [22] and Solovieff et al. [23] highlight some challenges regarding the study of pleiotropic traits in humans. One challenge is the computation of the statistical power and/or the minimum sample size necessary (MSSN) for genetic association, a critically important component of any gene mapping work. With these values, researchers may obtain a realistic estimate either of the MSSN to establish genetic associations or of the probability of detecting genetic associations for a collected sample. Power and MSSN calculations for single-phenotype tests of genetic association have been derived by Mitra [24] for the χ^2 test of independence on alleles/genotypes and by several authors [25–28] for the linear trend test (LTT). From this point forward, we refer to the former and the latter test as the “genotype test” (since the data collected are genotypes on individuals) and the “LTT,” respectively.

There have been a number of publications documenting ways to detect and analyze pleiotropic data, most recently for genome-wide association studies [23, 29–58], and also reporting methods to determine power and/or MSSN for association mapping [43, 45, 46, 59–66]. If one broadens the search to allow for multiple phenotypes that may not be pleiotropic, the list of published methods increases [34, 67–85]. Studying these methods, we noted that the majority deal with data analysis. We comment that a number of authors who document the power for their method do so by simulation [e.g., 45, 47] or for a specific data set [e.g., 6, 60, 86].

The purpose of this work was the development of an analytic approach to computing (1) the statistical power for a fixed sample size and a given significance level or (2) the MSSN (in terms of affected and unaffected individuals) to achieve a fixed power at a given significance level for a number of different statistical tests. Our method is threshold based, in the sense that we transform individuals with quantitative phenotype vector values into either affected or unaffected individuals using thresholds. From this point forward, we will use the abbreviations QT for “quantitative trait”/“quantitative phenotype” and QTV for “quantitative trait value” to refer to an individual’s quantitative phenotype vector values.

Our method is a natural extension of the univariate threshold-selected QT association power and the MSSN calculator [e.g., 87, 88], in that when the number of phenotypes is 1, our method is reduced to the univariate method. Some suggested benefits of our method are that (a) it is based on classic quantitative genetic mapping methods for selected sampling and (b) the mathematics used is well established and straightforward to implement.

We use a threshold approach because a number of pleiotropic diseases are defined this way. For example, Marfan syndrome and Tourette syndrome are composed of multiple traits, each of which may be caused by a single gene on the chromosome [2]. The phenotypes caused by these disorders are also quantitative or continuously distributed. That is, individuals may exhibit these traits to varying degrees (e.g., mild to severe). We note that each trait may be defined by thresholds for different QTs. Thresholds are provided below.

For the syndromes listed below, each of the conditions listed is necessary.

(1) Marfan syndrome: according to the Marfan Foundation [89], one definition of Marfan syndrome in the absence of a family history [90] encompasses (a) an aortic root dilatation Z score ≥ 2 and (b) a systemic score ≥ 7 points.

(2) For a person to be diagnosed with Tourette syndrome [91], he or she must (a) have ≥ 2 motor tics (e.g., blinking or shrugging shoulders), (b) have ≥ 1 vocal tic (e.g., humming, clearing the throat, or yelling out words or phrases), although they might not always happen at the same time, (c) have had tics (a) and (b) for ≥ 1 year (the tics can occur many times a day [usually in bouts] nearly every day, or on and off), (d) have tics that had started at ≤ 18 years of age, and (e) have symptoms that are not due to taking medicine or other drugs or due to having another medical condition (e.g., seizures, Huntington disease, or postviral encephalitis).

Additionally, we make a distinction between pleiotropy and locus heterogeneity. In Tourette syndrome, there is documented evidence of locus heterogeneity [92, 93]. Hence, in a particular family, it may be that these traits are “caused by a single gene” with a high penetrance. However, this situation is not what we mean by pleiotropy. For pleiotropy, it must be the same gene causing changes in multiple phenotypes across families/individuals.

We include a section on derivation of the power/MSSN for multivariate ANOVA (MANOVA) using the Pillai trace statistic applied to the quantitative measures directly. Our reasons are the following: (1) several published methods consider the power and/or MSSN for pleiotropic phenotypes using quantitative measures [31, 32, 36, 40, 42, 45, 94–101]; (2) while there is no uniformly most powerful test for MANOVA using equality of means as the null hypothesis, the Pillai trace statistic has high power in a number of different settings; and (3) the Pillai trace statistic is robust to several violations of assumptions in the MANOVA model [102, 103]. We perform a comparison of the MSSN for the Pillai statistic and our statistics using specified genetic model parameter settings.

Finally, we develop software that performs power and/or MSSN calculations for detecting genetic associations with (1) the LTT and the genotype test for threshold-defined phenotypes and (2) Pillai’s trace statistic for the original phenotypes. We note that this software is an extension of software programs designed to compute power and/or MSSNs considering a single locus and a single phenotype. In this work, MSSN calculations are for 2 traits (bivariate distributions) only. Our calculations may be extended to address any number of traits.

Methods¹

Test Statistic for One-Way MANOVA

Here, we present the test statistic used to test our multiple null hypotheses when the data are quantitative. Several multivariate mean vectors in a one-way MANOVA may be statistically compared using Wilks’s lambda, Pillai’s trace, Roy’s largest root, or Hotelling-Lawley’s tests [102, 103]. Though none of the tests is uniformly most powerful, Pillai’s trace statistic is reported to have good power in many scenarios and is robust to deviations from assumptions specified in MANOVA [102]. As an indication of its popularity, Pillai’s trace test is the default test in the *manova* function of the R statistical software package [106]. Wilks’s lambda is equivalent to the likelihood ratio test, and it has similar power to Pillai’s statistic in many alternative settings [102, 103].

Notation for the Pillai Statistic

Here, we define the null hypotheses for the Pillai statistic and the statistic itself:

g : Number of groups considered for each phenotype; here, g is the number of genotypes at a SNP locus, so that $g = 3$;
 p : Number of phenotypes (response variables).

Definition of the Pillai Statistic

Here, we present the Pillai trace statistic. It is used to test our multiple null hypotheses when the data are quantitative. As an indication of its popularity, Pillai’s trace test is the default test in the *manova* function of the R statistical software package [106]. Wilks’s lambda is equivalent to the likelihood ratio test and has similar power to Pillai’s statistic in many alternative settings [102, 103].

To begin, let

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_g \end{pmatrix} = \text{An } N \times p \text{ data matrix,}$$

where

$$Y_i = \begin{pmatrix} y_{i1} & y_{i2} & \cdots & y_{ip} \\ y_{i21} & y_{i22} & \cdots & y_{i2p} \\ \vdots & \vdots & & \vdots \\ y_{in_i1} & y_{in_i2} & \cdots & y_{in_ip} \end{pmatrix}$$

is an $n_i \times p$ data matrix, and y_{ijk} is the j -th observation of the i -th phenotype in the k -th genotype group, the total number of observations being denoted by $N = n_1 + \dots + n_g$. Note that $1 \leq i \leq g$, $1 \leq j \leq n_i$ for the i -th genotype group, and $1 \leq k \leq p$. Also, n_i is the number of individuals with the i -th genotype.

Let X denote the $N \times g$ design matrix given by

$$X = \begin{pmatrix} 1_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1_{n_g} \end{pmatrix},$$

where the matrices 1_{n_i} , $1 \leq i \leq g$, are of size $n_i \times 1$ and are defined as

$$1_{n_i} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Also, let $X'X$ and $1/N X'X$ be the diagonal $g \times g$ matrices given by

$$\begin{pmatrix} n_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & n_g \end{pmatrix}$$

and

$$\begin{pmatrix} \frac{n_1}{N} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{n_g}{N} \end{pmatrix},$$

¹ Notation for much of this section may be found in the Appendix.

respectively.

The Pillai trace test statistic is defined as

$$V = \sum_{i=1}^s \frac{\varphi_i}{1 + \varphi_i},$$

and is based on the $s = \min(g-1, p)$ eigenvalues $\{\varphi_1 \geq \dots \geq \varphi_s\}$ of $E^{-1}H$, where

$$E = A'(Y - X\hat{B})(Y - X\hat{B})A,$$

$$H = N(C\hat{B}A)'(C(1/N X'X)^{-1}C')^{-1}(C\hat{B}A).$$

Note that the matrix $\hat{B}$ is the matrix B with parameters estimated from the data. The matrices C and A are stated below. The estimate of each μ_{ij} is given by

$$\hat{\mu}_{ij} = \frac{1}{n_i} \sum_{u=1}^{n_i} y_{iuj}.$$

The Pillai statistic has an F distribution with $df_1 = r_C r_A$ and $df_2 = s(N - r_X + s - r_A)$ degrees of freedom under the null hypothesis. Note that r_C , r_A , and r_X are the ranks of the matrices C , A , and X , respectively.

Null Hypothesis

We can write a linear hypothesis in a one-way MANOVA as

$$H_0: CBA - D_0 = 0,$$

where

$$B = \begin{pmatrix} \mu_{11} & \dots & \mu_{1p} \\ \vdots & \ddots & \vdots \\ \mu_{g1} & \dots & \mu_{gp} \end{pmatrix}$$

is a $g \times p$ matrix for the p mean vectors. The matrices C and A are determined from a linear null hypothesis.

Power and Sample Size Calculations

O'Brien and Shieh [107] summarize the calculation of the power for global effects in one-way MANOVAs. The Pillai trace statistic under the alternative hypothesis has a noncentral F distribution with df_1 and df_2 degrees of freedom and the noncentrality parameter (NCP)

$$\lambda = Ns \left(\frac{V^*}{s - V^*} \right),$$

where

$$V^* = \sum_{i=1}^s \frac{\varphi_i^*}{1 + \varphi_i^*}$$

and φ_i^* is the i -th largest eigenvalue of

$$(A'SA)^{-1}(CBA - D_0)'(C(diag(p_1, \dots, p_g)^{-1}C')^{-1}(CBA - D_0),$$

where $p_j = n_j/N$ or the limit of the ratio as $N \rightarrow \infty$. We specify that the phenotype vectors in all groups have the common covariance matrix Σ . This common covariance matrix specification is necessary to derive the NCP. Note that for threshold-based phenotypes, we need not make such an assumption.

Example NCP Calculation for 2 Phenotypes

Consider our null hypothesis $H_0: \mu_{01} = \mu_{11} = \mu_{21}, \mu_{02} = \mu_{12} = \mu_{22}$ for 3 genotype groups ($i = 0, 1, 2$) with the bivariate phenotypes ($j = 1, 2$), that is, $p = 2$ and $g = 3$. Thus, $s = \min(g-1, p) = 2$. These

means are determined using the information from the section above (Methods, notation for QTs).

Let A be the 2×2 identity matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and let

$$C = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}.$$

Let

$$D_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

and let the covariance matrix of the bivariate phenotypes be denoted by

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}$$

with the correlation coefficient ρ . These matrices are specified so that we may test the null hypothesis $H_0: \mu_{01} = \mu_{11} = \mu_{21}, \mu_{02} = \mu_{12} = \mu_{22}$ stated above. We can calculate the 2×2 matrix Φ^* as

$$\Phi^* = (A'SA)^{-1}(CBA - D_0)'(C(diag(p_1, \dots, p_g))^{-1}C')^{-1}(CBA - D_0),$$

$$= \Sigma^{-1}(CB)' \left(C \left(diag \left(\frac{1}{p_0}, \frac{1}{p_1}, \frac{1}{p_2} \right) \right) C' \right)^{-1} (CB).$$

The matrix Φ^* is used to compute the eigenvalues φ_i^* which in turn are used to compute the Pillai statistic and the NCP.

Let us define the terms S_{ij} as

$$S_{ij} = \frac{1}{1 - \rho^2} \sum_{k=0}^2 p_k \left(\frac{\mu_{ki} - \bar{\mu}_i}{\sigma_i} \right) \left(\frac{\mu_{kj} - \bar{\mu}_j}{\sigma_j} \right),$$

where

$$\bar{\mu}_i = \sum_{k=0}^2 p_k \mu_{ki}.$$

We can simplify the matrix Φ^* to be:

$$\Phi^* = \begin{pmatrix} S_{11} - \rho S_{12} & \frac{\sigma_2}{\sigma_1} (S_{12} - \rho S_{22}) \\ \frac{\sigma_1}{\sigma_2} (S_{12} - \rho S_{11}) & S_{22} - \rho S_{11} \end{pmatrix}.$$

Note that

$$V^* = \sum_{i=1}^2 \frac{\varphi_i^*}{1 + \varphi_i^*} = \frac{\varphi_1^* + \varphi_2^* + 2\varphi_1^*\varphi_2^*}{1 + \varphi_1^* + \varphi_2^* + \varphi_1^*\varphi_2^*},$$

$$\varphi_1^* + \varphi_2^* = \text{trace}(\Phi^*) = S_{11} + S_{22} - 2\rho S_{12},$$

and

$$\varphi_1^* \varphi_2^* = \det(\Phi^*),$$

$$= (S_{11} - \rho S_{12})(S_{22} - \rho S_{11}) - \frac{\sigma_2}{\sigma_1} (S_{12} - \rho S_{22}) \frac{\sigma_1}{\sigma_2} (S_{12} - \rho S_{11}),$$

$$= (1 - \rho^2)(S_{11}S_{22} - S_{12}^2).$$

Therefore, the NCP λ can be written as

$$\begin{aligned}\lambda &= N \frac{sV^*}{s - V^*}, \\ &= N \frac{s \times \text{trace}(\Phi^*) + 2s \times \det(\Phi^*)}{s + (s-1) \times \text{trace}(\Phi^*) + (s-2) \times \det(\Phi^*)}, \\ &= N \frac{2(S_{11} + S_{22} - 2\rho S_{12}) + 4(1 - \rho^2)(S_{11}S_{22} - S_{12}^2)}{2 + S_{11} + S_{22} - 2\rho S_{12}}.\end{aligned}\quad (1)$$

The power of the Pillai trace test is obtained by

$$\Pr(F(df_1, df_2, \lambda) \geq f_{\alpha, df_1, df_2}),$$

where f_{α, df_1, df_2} is the $(1 - \alpha)$ quantile of a central F distribution with df_1 and df_2 degrees of freedom, respectively, and $F(df_1, df_2, \lambda)$ is a noncentral F random variable with NCP λ and degrees of freedom df_1 and df_2 , respectively. For our example, $df_1 = r_{CA} = 4$ and $df_2 = s(N - r_X + s - r_A) = 2(N - 3)$.

Bivariate Example

For the remainder of this work (excluding the Discussion), we focus on bivariate distribution, that is, on pleiotropic diseases with 2 QTs. We do this because results are more easily interpreted, and because we can present graphs of functions such as the cumulative distribution function.

MSSN Calculations Using a Factorial Design

We asked the following question: which factors most substantially alter the calculated MSSN when testing for genetic associations with a pleiotropic gene affecting 2 phenotypes?

To answer this question, we used a $2^4 \times 3^3$ factorial design [see 108] on a total of 7 design variables (factors) to approximate the calculated MSSN with functions of the design variables. These factors are listed in Table 1. Note that we obtained $2^4 \times 3^3 = 432$ vectors of factor settings and therefore 432 MSSN calculations. One benefit of the factorial design is that we can look at multiple factors jointly over a broad range of settings and assess the factors that change the outcome variable the most. For all MSSN calculations, we specified that the fixed power is 0.80 and the significance level is 5×10^{-8} .

Approximation of the Calculated MSSN

After we computed all 216 MSSN values for the Pillai test, as well as all 432 MSSN values (we compute the number of affected individuals needed and set the number of unaffected individuals to be equal to the number of affected individuals, i.e., $r = 1$) for the genotype test and the LTT, we performed a linear model analysis (i.e., ANOVA) on the 7 main factors (Table 1) and all 2-way interactions. The ANOVA calculations were performed using the methods developed for the R statistical software package [106].

Our rationale for performing the ANOVA with the factorial design was as follows: Equation 1 above and Equations A8.1 and A9.1 in the online supplementary material (for all online suppl. material, see www.karger.com/doi/10.1159/000457135) are closed-form equations that specify the NCPs (from which the MSSN may be calculated). Here, the MSSN is given by $n = n(r, w_k, g_{ik})$, where i = affection status, k = genotype. Although they are analytic, it is difficult to identify the variables that are most impor-

Table 1. Factors (and their settings) used in the MSSN calculations for genetic tests of association

Factor	Number of settings	Setting values
p_d	2	0.05, 0.330
σ_1^2	2	0.05, 0.10
τ_1	3	-0.50, 0.00, 0.50
σ_2^2	2	0.025, 0.05
τ_2	3	-0.50, 0.00, 0.50
ρ	3	0.00, 0.33, 0.67
Percent-affected and percent-unaffected	2	10%, 25%

MSSN, minimum sample size necessary; p_d , disease allele frequency; σ_1^2 , variance for the first phenotype's quantitative trait distribution; τ_1 , dominance-additivity ratio for the first phenotype; σ_2^2 , variance for the second phenotype's quantitative trait distribution; τ_2 , dominance-additivity ratio for the second phenotype; ρ , correlation between the 2 phenotypes, or ρ_{12} . While we can consider negative correlations, for bivariate distributions, 2 phenotypes may always be parameterized so that the correlation is non-negative.

tant. Consequently, we approximated the exact function by a linear model (including all 2-way interactions) $\hat{n}(r, w_k, g_{ik}) = \mu + \mu_r + \dots$. We used 432 settings for our linear model approximation (216 for the Pillai statistic, since it is not dependent upon percent-affected and percent-unaffected settings) and report the factors that most fully explain the MSSN.

We note here and in the Results section that we do not attempt to make statistical inferences from our applications of the factorial design and ANOVA. Rather, we use them as explanatory tools specifically documenting the factors (main and interaction) that appear to have the most substantial effect on altering the MSSN (i.e., those with the largest F -statistics), and then documenting quantitatively whether the results appear to be true. We can do this by computing MSSNs considering different settings of the aforementioned factors and checking whether the different settings produce substantially different MSSN estimates.

Results

Factors that Most Significantly Alter the Genetic Association Test MSSN

Genotype Test

In Table 2, we report the results of our ANOVA for the genotype test. Overall, this statistic on average had the smallest MSSN requirements for any set of factor settings in Table 1. This result is notable, since the genotype test has 2 degrees of freedom (df); thus, one might expect the LTT to have lower MSSN values. Also, the genotype test

Table 2. Results of the analysis of variance for main effects and all 2-way interactions (genotype test)

Factor	df	SSQ _{Factor}	F-statistic	φ^2
Percent-affected	1	1,560,721	33,815.121	0.314
ρ	2	1,723,434	18,670.263	0.347
σ_1^2	1	612,234	13,264.884	0.123
σ_2^2	1	308,685	6,688.076	0.062
p_d	1	303,127	6,567.645	0.061
$\rho \times$ percent-affected	2	103,543	1,121.697	0.021
$\sigma_1^2 \times$ percent-affected	1	46,967	1,017.613	0.009
$p_d \times$ percent-affected	1	40,969	887.657	0.008
$\sigma_1^2 \times \rho$	2	63,336	686.134	0.013
$\sigma_1^2 \times \sigma_2^2$	1	25,551	553.597	0.005
τ_1	2	46,357	502.194	0.009
$\sigma_2^2 \times$ percent-affected	1	22,923	496.648	0.005
$\sigma_2^2 \times \rho$	2	31,059	336.47	0.006
$p_d \times \rho$	2	23,991	259.896	0.005
τ_2	2	16,522	178.984	0.003
$\tau_1 \times$ percent-affected	2	5,191	56.235	0.001
$p_d \times \sigma_2^2$	1	2,162	46.851	0
$\tau_2 \times$ percent-affected	2	2,892	31.331	0.001
$\tau_1 \times \rho$	4	4,434	24.018	0.001
$\tau_1 \times \sigma_2^2$	2	1,723	18.67	0
$\tau_1 \times \tau_2$	4	3,101	16.796	0.001
$\sigma_1^2 \times \tau_2$	2	1,041	11.28	0
$\tau_2 \times \rho$	4	1,606	8.7	0
$p_d \times \sigma_1^2$	1	282	6.11	0
$\sigma_2^2 \times \tau_2$	2	97	1.051	0
$\sigma_1^2 \times \tau_1$	2	74	0.799	0
$p_d \times \tau_1$	2	70	0.753	0
$p_d \times \tau_2$	2	2	0.02	0
Residuals	379	17,493		
Total		4,964,426		

The values in the column labeled “Factor” are defined in Table 1. The column SSQ_{Factor} is the sum of squares for the given factor. The column labeled “ φ^2 ” lists each factor’s proportion of the overall sum of squares. That is, $\varphi^2 = \text{SSQ}_{\text{Factor}} / \text{SSQ}_{\text{Total}}$. All values with exception of those in the last column are computed using methods developed for the R statistical software package [106].

is applied to categorical data, and it is generally true that for quantitative data, quantitative data-based tests such as Pillai’s will require smaller MSSNs than do tests on categorical data. We examine this point further in the Discussion section.

In Table 2, the factors are sorted from the largest to the smallest F -statistic. Also, we report the value φ^2 , the respective factor’s proportion of the overall sum of squares (SSQ). Specifically,

$$\varphi^2 = \frac{\text{SSQ}_{\text{Factor}}}{\text{SSQ}_{\text{Total}}}$$

(values are provided in Table 2). Based on the F -statistics and the φ^2 values, we may infer that there are 5 main factors that most substantially influence the number of affected individuals needed to detect an association. These are, in order of the F -statistic (rounded to nearest integer from Table 2): percent-affected (F -statistic = 33,815); ρ (correlation) (F -statistic = 18,670); σ_1^2 (F -statistic = 13,265); σ_2^2 (F -statistic = 6,688); and p_d (F -statistic = 6,568). Along with their 2-way interaction terms (a total of 10), these 5 factors account for 98% of the proportion of the total SSQ (SSQ_{Total}) (Table 2). The dominance-additivity ratios τ_1 and τ_2 had a relatively small impact on the calculated MSSN. This result suggests that the genotype test is equally powerful when the QT loci (QTLs) operate in either an additive or a nonadditive mode of inheritance. That is, researchers need not focus on whether their traits of interest deviate from an additive mode of inheritance when performing MSSN calculations.

Given these results, we performed a regression analysis in which we used the 5 main-effect terms and their 2-way interaction. The results of the regression analysis are provided in Table 3. As main be seen in Table 3 and Equation 2 below, there are actually 6 “main”-effect terms, since there are 3 settings for the correlation factor ρ ; hence, we need 2 separate variables. Our goal was to compute the coefficients of the fitted sample size equation:

$$\hat{n}_A = \beta_0 + \sum_{d=1}^D \sum_{i=1}^{\psi_d} \beta_i x_i + \sum_{d=1}^D \sum_{f=2}^D \sum_{i=1}^{\psi_d} \sum_{j=1}^{\psi_f} \beta_{ij} x_i x_j + e, \text{ where } e \sim N(0, \epsilon^2). \quad (2)$$

Here, D is the number of factors (5 in this case), and ψ_z is the number of df for the z -th factor, $1 \leq z \leq D$. Also, $1 \leq d < f \leq D$, and $\beta_{ij} = 0$ if i, j are settings for the same factor. This form of the fitted equation is used for all test statistics (genotype, LTT, and Pillai).

From Table 3, we compute the fitted function as

$$\hat{n} = 154.718 + 139.272x_1 + 81.701x_2 + 185.045x_3 - 43.27x_4 - 38.942x_5 - 21.689x_6 + 31.973x_1x_2 + 75.548x_1x_3 - 41.708x_1x_4 - 29.137x_1x_5 - 38.954x_1x_6 + 0.00x_2x_3 - 25.374x_2x_4 - 18.002x_2x_5 - 17.221x_2x_6 - 59.121x_3x_4 - 41.421x_3x_5 - 36.489x_3x_6 + 30.763x_4x_5 + 3.232x_4x_6 + 8.949x_5x_6, \quad (3)$$

where:

$$x_1 = \begin{cases} 1, & \text{if percent - affected} = 25\% \\ 0, & \text{if percent - affected} = 10\% \end{cases},$$

$$x_2 = \begin{cases} 1, & \rho = 0.33 \\ 0, & \rho = \text{otherwise} \end{cases},$$

$$x_3 = \begin{cases} 1, & \rho = 0.67 \\ 0, & \rho = \text{otherwise} \end{cases},$$

Table 3. Coefficients for the linear regression model using the 5 most significant main factors (genotype test)

Factor and setting	Coefficient estimate	Standard error	<i>t</i> statistic
(Intercept)	154.718	3.449	44.853
Percent-affected = 25	139.272	3.688	37.767
$\rho = 0.33$	81.701	4.123	19.816
$\rho = 0.67$	185.045	4.123	44.882
$\sigma_1^2 = 0.10$	-43.27	3.688	-11.734
$\sigma_2^2 = 0.05$	-38.942	3.688	-10.56
$p_d = 0.33$	-21.689	3.688	-5.882
Percent-affected = 25, $\rho = 0.33$	31.973	3.688	8.67
Percent-affected = 25, $\rho = 0.67$	75.548	3.688	20.487
Percent-affected = 25, $\sigma_1^2 = 0.10$	-41.708	3.011	-13.852
Percent-affected = 25, $\sigma_2^2 = 0.05$	-29.137	3.011	-9.677
Percent-affected = 25, $p_d = 0.33$	-38.954	3.011	-12.937
$\rho = 0.33$, $\sigma_1^2 = 0.10$	-25.374	3.688	-6.881
$\rho = 0.33$, $\sigma_2^2 = 0.05$	-18.002	3.688	-4.882
$\rho = 0.33$, $p_d = 0.33$	-17.221	3.688	-4.67
$\rho = 0.67$, $\sigma_1^2 = 0.10$	-59.121	3.688	-16.032
$\rho = 0.67$, $\sigma_2^2 = 0.05$	-41.421	3.688	-11.233
$\rho = 0.67$, $p_d = 0.33$	-36.489	3.688	-9.895
$\sigma_1^2 = 0.10$, $\sigma_2^2 = 0.05$	30.763	3.011	10.217
$\sigma_1^2 = 0.10$, $p_d = 0.33$	3.232	3.011	1.073
$\sigma_2^2 = 0.05$, $p_d = 0.33$	8.949	3.011	2.972

Here, we present the results of a linear regression using the 5 most significant factors from Table 2. We include all 2-way interactions of these factors. An example description of the factors is as follows: “ $\rho = 0.33$ ” means: if the setting of correlation is 0.33, use the coefficient 81.701 (second column) when computing the fitted value. Otherwise, use 0. We computed coefficients for the other main factors in the same manner. For the 2-way interactions, consider the example “percent-affected = 25, $p_d = 0.33$.” Here, if the disease allele frequency setting is 0.33 and the percent-affected setting is 25, then the coefficient used for the fitted values is -38.954, otherwise it is 0. All values were computed using methods (specifically, the *lm* and *summary* commands) developed for the R statistical software package [106]. All values in the last 3 columns are rounded to 3 decimal places.

$$x_4 = \begin{cases} 1, \sigma_1^2 = 0.10 \\ 0, \sigma_1^2 = 0.05 \end{cases},$$

$$x_5 = \begin{cases} 1, \sigma_2^2 = 0.050 \\ 0, \sigma_2^2 = 0.025 \end{cases},$$

$$x_6 = \begin{cases} 1, p_d = 0.33 \\ 0, p_d = 0.05 \end{cases}.$$

Reviewing the coefficients in Equation 2, we observe that increasing the percent-affected from 10 to 25% produces a substantial increase in MSSN (approx. 139 individuals; coefficient for variable x_1). The next-largest coefficient is for the correlation term ρ in the variance-covariance matrix Σ . Increasing the correlation from 0 (uncorrelated phenotypes) to 0.33 produces an increase in MSSN of approximately 82 individuals (coefficient for variable x_2), and increasing the correlation from 0 to 0.67 produces an increase in MSSN of 185. This coefficient is

the single largest coefficient in the fitted Equation 1. Coefficients for the other main effects are smaller, but significantly nonzero.

For the interaction terms, the larger coefficient in Equation 3 in absolute values is for the pair (percent-affected, ρ). When percent-affected equals 25 and ρ equals 0.67, the increase in MSSN is approximately 76. With the exception of the pairs (σ_1^2, p_d) and (σ_2^2, p_d) , the coefficients for all the other interaction terms are >15 in absolute values (Equation 2; Table 3). These results are consistent with the *F*-statistic values in Table 2.

Finally, a review of the results in Table 3 suggests that the MSSN is decreased the most when $\sigma_1^2 = 0.10$, since every coefficient that contains $\sigma_1^2 = 0.10$ (with the exception of coefficients for the third-to-last and second-to-last rows of Table 3) is negative. This result is consistent with the fact that increasing QTL variance increases the separation among the component multivariate normal distri-

Fig. 1. Scatter plot of the fitted minimum sample size necessary (MSSN) versus the analytic MSSN for the genotype test for 432 factor settings. Each triangle represents the coordinates (genotype test fitted MSSN based on Equation 3, genotype test analytic MSSN). The equation in this figure is the linear trend line equation as computed using Microsoft Excel. MSSNs were computed using the vector of settings ($x_1, \dots, x_6$) (Equation 3). The significance level was 5×10^{-8} .

butions, thereby making it easier to determine genotypes from QTVs.

In Figure 1, we present a plot of the fitted values (using Equation 3) versus the analytic MSSN ($n = n_A + n_U$) determined using the NCP (online suppl. material, Equations A8.1 and A8.2). The coefficients of the trend line, computed using the method in Excel, are consistent with the finding that the analytic MSSNs are accurately approximated by a linear combination of the 6 variables $x_1, \dots, x_6$ (Table 3) and their 2-way interactions. We base this conclusion on the fact that the trend line intercept is 0.0005 (close to 0) and the slope is exactly 1. From this we may conclude that for the parameter settings considered in Table 1, only 5 of the 7 factors are needed to approximate the analytic MSSN, and that among them, percent-affected/unaffected and the correlation ρ make the greatest change. Since percent-affected/unaffected is the only variable that researchers can control, in order to decrease MSSN requirements, one should decrease the percent-affected value to a 10% threshold (set x_1 to 0 in Equation 3). Doing so will decrease the fitted MSSN by approximately 139 individuals (coefficient of x_1 in Equation 3). In the online supplementary material, we computed analytic MSSNs over a range of percent-affected/unaffected values for the genotype test and the LTT and document that as the percent-affected/unaffected setting approaches 0%, so does the MSSN (online suppl. material, Fig. A4).

Linear Trend Test

The results of the LTT are very similar to those of the genotype test, although the MSSN requirements are generally higher. We placed the results of our analyses in the online supplementary material (Table A2). Also, see the Discussion section.

Pillai Test

We provide the results of our ANOVA for the Pillai test in Table 4. Overall, this statistic had the largest MSSN requirements for any set of factor settings in Table 1. Note that the factor percent-affected/unaffected is not used when computing MSSN requirements for the Pillai statistic, because we use QTVs on all individuals, not just those whose values are above/below a threshold. Hence, we computed the ANOVA for a total of $432/2 = 216$ vectors of settings from Table 2.

As in Table 2, the factors considered in our ANOVA are sorted from the largest to the smallest F -statistic, and we report the ϕ^2 values (listed in Table 4). Considering the F -statistics and the ϕ^2 values, we infer that there are 3 main terms that most substantially affect the MSSN to detect associations. These are, in order of the F -statistic (rounded to nearest integer): σ_1^2 (F -statistic = 5,804); σ_2^2 (F -statistic = 630); and ρ (F -Statistic = 559). The three 2-order interactions of these terms are: $\sigma_1^2 \times \sigma_2^2$ (F -statistic = 297), $\sigma_1^2 \times \rho$ (F -statistic = 155), and $\sigma_2^2 \times \rho$ (F -statistic = 14). These 6 main and interaction factors account for

Table 4. Results of the analysis of variance for main effects and all 2-way interactions (Pillai test)

Factor	df	SSQ _{Factor}	F-statistic	φ^2
σ_1^2	1	4,626,159	5,804.04	0.679
σ_2^2	1	502,017	629.836	0.074
ρ	2	891,480	559.231	0.131
$\sigma_1^2 \times \sigma_2^2$	1	237,057	297.415	0.035
$\sigma_1^2 \times \rho$	2	247,063	154.984	0.036
$\tau_1 \times \tau_2$	4	83,801	26.284	0.012
p_d	1	16,801	21.078	0.002
$\sigma_2^2 \times \rho$	2	22,746	14.269	0.003
$p_d \times \rho$	2	12,947	8.122	0.002
τ_1	2	9,030	5.665	0.001
τ_2	2	9,030	5.665	0.001
$p_d \times \sigma_1^2$	1	2,308	2.896	0
$\tau_1 \times \rho$	4	6,988	2.192	0.001
$\tau_2 \times \rho$	4	6,988	2.192	0.001
$p_d \times \tau_1$	2	1,403	0.88	0
$p_d \times \tau_2$	2	1,403	0.88	0
$\sigma_1^2 \times \tau_1$	2	1,243	0.78	0
$\sigma_1^2 \times \tau_2$	2	1,243	0.78	0
$p_d \times \sigma_2^2$	1	28	0.035	0
$\sigma_2^2 \times \tau_1$	2	16	0.01	0
$\sigma_2^2 \times \tau_2$	2	16	0.01	0
Residuals	173	137,891		
Total		6,817,658		

The legend to this table is virtually identical to the legend to Table 2, with the exception that the “percent-affected” factor is not considered, since the Pillai statistic is computed on all individuals. All values with the exception of those in the last column were computed using methods developed for the R statistical software package [106].

approximately 96% of the proportion of the SSQ_{Total} (Table 4, last column). These results suggest that a linear function of the top 5 factors (like Equation 3 for the genotype test) provides a very close approximation to the actual MSSN for all 216 vectors of settings from Table 1.

Using the results in Table 4, we performed a regression analysis in which we selected the 3 main-effect terms (a total of 4 variables, given the 2 settings of correlation) and their 2-way interactions. We present the results in Table 5.

From Table 5, we computed the fitted function as

$$\hat{n} = 651.081 - 277.541x_1 - 173.81x_2 + 150.831x_3 + 215.882x_4 + 132.513x_1x_2 - 78.614x_1x_3 - 165.614x_1x_4 - 6.512x_2x_3 + 39.915x_2x_4 \quad (4)$$

where

$$x_1 = \begin{cases} 1, \sigma_1^2 = 0.10 \\ 0, \sigma_1^2 = 0.05 \end{cases},$$

$$x_2 = \begin{cases} 1, \sigma_2^2 = 0.050 \\ 0, \sigma_2^2 = 0.025 \end{cases},$$

$$x_3 = \begin{cases} 1, \rho = 0.33 \\ 0, \rho = \text{otherwise} \end{cases},$$

$$x_4 = \begin{cases} 1, \rho = 0.67 \\ 0, \rho = \text{otherwise} \end{cases}.$$

Studying Equation 4, we note that changes in main factors result in changes of at least 174 individuals. For example, increasing σ_1^2 from 0.05 to 0.10 reduces the MSSN by 278 individuals in Equation 3. Similarly, increasing the correlation ρ from 0 to 0.33 increases the MSSN by 151. For the interaction terms, the largest change is –166, occurring when σ_1^2 is 0.10 and ρ is 0.67. The smallest change in MSSN occurs when σ_2^2 is 0.05 and ρ is 0.33.

In Figure 2, we plotted the fitted values (using Equation 4) versus the analytic MSSN ($n = n_A + n_U$) determined using the Pillai NCP (online suppl. material). As with Figure 1, the coefficients of the trend line, computed using the method in Excel, are consistent with the finding that the analytic MSSNs are accurately represented by a linear combination of all terms in Equation 4 (the trend line intercept is 0.0004, the slope is 1.0). In contrast to the genotype test results, for the Pillai test, we required only 3 of the 6 factors to approximate the analytic MSSN (Table 6; Fig. 3). Also, the MSSN requirements had decreased most substantially by increasing the QTL variances σ_1^2 and σ_2^2 and by decreasing the correlation ρ .

Which Method Produces the Smallest MSSN Requirements?

So far, we have answered the questions of which factors most substantially alter MSSN requirements, and by how much, for the genotype test, the Pillai test, and the LTT (online suppl. material) for the factor settings in Table 1. An equally important question is: which statistic produces the smallest analytic MSSN requirements for any vector of factor settings in Table 1? To answer this question, we computed the 5 sets of differences:

- I. LTT($p_d, \sigma_1^2, \tau_1, \dots, \rho$, percent-affected) – genotype($p_d, \sigma_1^2, \tau_1, \dots, \rho$, percent-affected);
- II. Genotype($p_d, \sigma_1^2, \tau_1, \dots, \rho$, percent-affected = 10) – Pillai($p_d, \sigma_1^2, \tau_1, \dots, \rho$, percent-affected = 10);
- III. Genotype($p_d, \sigma_1^2, \tau_1, \dots, \rho$, percent-affected = 25) – Pillai($p_d, \sigma_1^2, \tau_1, \dots, \rho$, percent-affected = 25);

Fig. 2. Scatter plot of the fitted minimum sample size necessary (MSSN) versus the analytic MSSN for the Pillai test using 216 vectors of factor settings. Each triangle represents the coordinates (Pillai test fitted MSSN based on Equation 4, Pillai test analytic MSSN). The explanations in the legend to Figure 1 apply to this figure as well.

- IV. LTT($pd, \sigma_1^2, \tau_1, \dots, \rho$, percent-affected = 10) – Pillai($pd, \sigma_1^2, \tau_1, \dots, \rho$, percent-affected = 10);
- V. LTT($pd, \sigma_1^2, \tau_1, \dots, \rho$, percent-affected = 25) – Pillai($pd, \sigma_1^2, \tau_1, \dots, \rho$, percent-affected = 25).

Each of the differences in MSSN is computed as a function of the parameter settings. For example, if $p_d = 0.33$, $\sigma_1^2 = 0.10$, $\tau_1 = 0.0$, $\sigma_2^2 = 0.05$, $\tau_2 = 0.5$, $\rho = 0.0$, and percent-affected = 25, then Difference I is:

Analytic MSSN for LTT for vector (0.33, 0.10, 0.0, 0.05, 0.5, 0.0, 25) – Analytic MSSN for genotype test for vector (0.33, 0.10, 0.0, 0.05, 0.5, 0.0, 25).

Differences II–V are computed with a fixed value for the last parameter (percent-affected). The reason is that the Pillai test is a function of only 6 parameters in Table 1; as noted previously, it is not a function of the percent-affected parameter. For each of the Differences I–V, we present the empirical distributions of the results in the form of box plots. These box plots may be found in Figure 3.

Note that Difference I is computed over 432 vectors, while Differences II–V are computed over 216 vectors. Some of the key findings resulting from a study of Figure 3 are that the genotype test usually has the smallest sample size (previously mentioned) and that the genotype test and the LTT almost always require smaller analytic MSSNs than does the Pillai test. In fact, viewing the 4 right-most box plots, the greatest difference between the Pillai and any of the other test statistics, where Pillai requires a

Table 5. Coefficients for the linear regression model using the 3 most significant main factors and all interactions (Pillai test)

Factor	Coefficient estimate	Standard error	<i>t</i> statistic
(Intercept)	651.081	8.089	80.491
$\sigma_1^2 = 0.10$	–277.541	10.232	–27.126
$\sigma_2^2 = 0.05$	–173.81	10.232	–16.987
$\rho = 0.33$	150.831	10.852	13.898
$\rho = 0.67$	215.882	10.852	19.893
$\sigma_1^2 = 0.10, \sigma_2^2 = 0.05$	132.513	10.232	12.951
$\sigma_1^2 = 0.10, \rho = 0.33$	–78.614	12.531	–6.273
$\sigma_1^2 = 0.10, \rho = 0.67$	–165.614	12.531	–13.216
$\sigma_2^2 = 0.05, \rho = 0.33$	–6.512	12.531	–0.52
$\sigma_2^2 = 0.05, \rho = 0.67$	39.915	12.531	3.185

In this table, we present the linear regression analysis coefficients for the 3 most significant factors from Table 4. Also, we include all 2-way interaction terms. Similar to Table 3, we have the following factor descriptions: “ $\sigma_1^2 = 0.10$ ” means: if the setting of the first phenotype’s quantitative trait locus variance is 0.10, use the coefficient –277.541 (second column) when computing the fitted value. Otherwise, use 0. We computed coefficients for the other main factors in the same manner. Computation for the interaction factors is described in the legend to Table 3. All values were computed using methods (specifically, the *lm* and *summary* commands) developed for the R statistical software package [106].

Fig. 3. Box plots for all pairs of statistical test differences in analytic MSSN. $\diamond$, mean value of differences; upper horizontal end of gray box, 3rd quartile (3Q) of values (75% of the differences are less than the value corresponding to this line); black horizontal line inside gray box, median value (50% of the differences are less than the value corresponding to this line and 50% are greater than the value); lower horizontal end of gray box, 1st quartile (1Q) of values (75% of the differences are greater than the value corresponding to this line); end of upper whisker, maximum value for the set of differences x that satisfy the condition $1Q - 1.5\delta \leq x \leq 1.5\delta + 3Q$, $\delta = 3Q - 1Q =$ interquartile range; end of lower whisker, minimum value for the set of differences x that satisfy the inequality listed directly above; *, value y that satisfies either $1.5\delta + 3Q < y \leq 3\delta + 3Q$ or $1Q - 3\delta \leq y < 1Q - 1.5\delta$; $\circ$, outlier, value z that satisfies either $3\delta + 3Q < z$ or $1Q - 3\delta > z$.

Table 6. Percentiles for MSSN ratios with different test statistics

Percentile	Ratio of MSSNs				
	LTT/ genotype	Pillai/ genotype (10%)	Pillai/ genotype (25%)	Pillai/ LTT (10%)	Pillai/ LTT (25%)
Minimum	0.95	1.59	0.94	1.20	0.74
Median	1.35	3.41	1.95	2.62	1.65
Mean	1.26	3.37	1.98	2.66	1.64
Maximum	1.64	5.28	3.14	4.18	2.45

In this table, we use the abbreviations “LTT ($x\%$)” and “genotype ($x\%$)” to signify the MSSNs for the LTT and the genotype test, respectively, when the percent-affected/unaffected settings are x ($x = 10$ or 25%). Also, each column’s pair of tests corresponds to the same numbered column in Figure 3. For example, the first pair of tests is the LTT and the genotype test. The same pair is considered in the first column of Figure 3. MSSN, minimum sample size necessary.

smaller sample size, is for the “LTT (25%) – Pillai” box plot (the right-most one in Fig. 3). The difference is 124 (outlier for “LTT (25%) – Pillai”; Fig. 3). In results not shown, this difference occurs for the vector of settings $p_d = 0.33$, $\sigma_1^2 = 0.10$, $\tau_1 = -0.50$, $\sigma_2^2 = 0.025$, $\tau_2 = 0.50$, $\rho = 0.67$, percent-affected = 25. For this vector, the LTT analytic MSSN is 477 and the Pillai test analytic MSSN is 353.

In Table 6, we present the differences in Figure 3 as ratios. Lehmann and Romano [109], among others, defined these ratios as asymptotic relative efficiencies. We report the minimum, median, mean, and maximum ratios for all pairs of test statistics. In this way, we could

compare results across columns. The smallest median and mean values – 1.35 and 1.26, respectively – were for the LTT/genotype MSSN ratio. This result suggests that the MSSNs for these 2 test statistics are most similar. The largest median and mean values of 3.41 and 3.37 were for the Pillai/genotype (10%) MSSN ratio. This result is consistent with the fact that the “genotype (10%) – Pillai” MSSN box plot has the lowest range of differences (vertical axis) in Figure 3.

For all ratios below the median ratio of 1.35 for the LTT/genotype MSSN ratio, every vector has the disease allele frequency setting $p_d = 0.05$. This result suggests that

LTT and genotype test MSSNs are most similar for smaller disease allele frequencies.

Finally, we note that we have developed software to perform these calculations. This software will be made available online within the near future. Researchers who want stand-alone copies of the software may contact the first author.

Discussion

In this work, we presented the method (the genotype test) for computing asymptotic power and MSSN calculations for genetic associations with pleiotropic traits. In our design, affection status is defined through thresholds. We included computations of power and MSSN for MANOVA by applying Pillai's statistic.

The first observation we make is that we could specify a multivariate function to compute probabilities for pleiotropic phenotypes (Formulas A1 and A2 in the online suppl. material). Also, we derived categorical data from the QTVs and applied the genotype test and LTT to the categorical data (Equation A4 in the online suppl. material). Furthermore, we computed analytic power and MSSN formulas for the genotype test and LTT (Formulas A8.1 and A9.1 in the online suppl. material), as well as analytic power and MSSN formulas for the Pillai MANOVA test applied to all QTVs.

Our ANOVA results for the factorial designs indicate that, for the genotype test, the factors that most substantially alter MSSNs are correlations between the 2 QTs (ρ) and the percent-affected/unaffected settings. From the results from Table 3 and Equation 3, we see that the MSSN decreases with a decrease in the correlation and a change of the percent-affected/unaffected setting from 25 to 10%. Changes in these 2 factors reduce the MSSN for the LTT as well (results not shown). We comment that we used the ANOVA to provide a numerical approximation (with linear and 2-way interaction terms) to the analytic formulas for the MSSN. The factors we considered in the approximation are those with the largest F -statistic values.

For the Pillai test, the analytic MSSN is accurately described by settings in 3 factors and their interactions: σ_1^2 , σ_2^2 , and ρ (Table 5; Equation 4). Increases in the QTL variances σ_1^2 , σ_2^2 reduced the MSSN, while a decrease in the correlation ρ produced a decrease in the MSSN.

When comparing all the MSSNs for all tests, we see that the genotype test usually requires the smallest MSSN to achieve 80% power at the 5×10^{-8} significance level for the vector of settings in Table 1. We draw this conclusion

by studying the box plots of MSSN differences for all pairs of test statistics. The only test statistic that has a smaller MSSN than the genotype test for any significant portion of vector settings is the LTT. In fact, for 110/432 (25%) of the vectors, the LTT has an MSSN that is as small as or smaller than that of the genotype test. However, the maximum difference is 14 individuals, and the relative efficiency is never less than 95% (Table 6).

While this work focused on sample size calculations, through use of NCPs we can just as easily perform power calculations for a fixed sample size. The conclusions we draw about the 3 statistics are the same (e.g., the genotype test has the largest power on average for the different vectors of factor settings, followed by the LTT, etc.) (data not shown).

What if a SNP we are studying is in linkage disequilibrium with a disease gene but not the gene itself [23]? In such circumstances, we use the method implemented by others [e.g., 87, 88] to perform power and MSSN calculations of threshold-selected QTLs that are in linkage disequilibrium with a disease locus.

A final and very important issue to address is the fact that the Pillai test, which is applied to quantitative data for all individuals, has larger MSSN values than either the genotype test or the LTT. Our explanation for this result is that our design focuses on MSSN calculations *before* any data are collected. Also, our focus is on gene mapping, not on tests of linearity. If one were conducting a population-based study, where phenotype and genotype values were collected on all individuals, and all 3 test statistics were applied to all individuals, then the Pillai statistic would typically have the smallest sample size requirement.

Consider the following example of vector settings: $p_d = 0.05$, $\sigma_1^2 = 0.10$, $\tau_1 = 0.0$, $\sigma_2^2 = 0.05$, $\tau_2 = 0.50$, $\rho = 0.0$, percent-affected-phenotype 01 = (top) 100%, percent-affected-phenotype 02 = (top) 50%, percent-unaffected-phenotype 01 = (lower) 100%, percent-unaffected-phenotype 02 = (lower) 50%. The parameter settings (with the exception of percent-affected and percent-unaffected) are taken from Table 1.

Regarding the affection thresholds, imagine a square. If we draw a horizontal line through the square, cutting it in half, affected individuals are those subjects whose pair of QTVs are in the upper half of the square, and unaffected individuals are those subjects whose pair of QTVs are in the lower half. With these thresholds, we use all the individuals for the genotype test and LTT, as well as the Pillai test. Applying our formulas, we compute that MSSNs are 1,471 for the genotype test, 1,387 for the LTT,

and 326 for the Pillai test for a 5×10^{-8} significance level. The Pillai MSSN is much lower than that for either of the categorical data-based tests.

Similarly, if we define affection by using a vertical line rather than a horizontal line, our MSSNs are 836 for the genotype test, 785 for the LTT, and 326 for the Pillai test (the Pillai statistic is not dependent upon threshold settings). That is, the Pillai MSSN is less than half of that of either of the categorical data-based tests.

Another practical issue regarding lower values for percent-affected (like 10%) is that for small or moderate MSSNs, one may not observe individuals with phenotypes in this region. For small and moderate MSSNs, the thresholds may be theoretically desirable but impractical. In such circumstances, one might have no choice but to increase the percent-affected threshold.

Finally, we comment that the software to perform power and sample size calculations for pleiotropy is freely available for Windows and Ubuntu Linux. We anticipate to have a Web-based and/or R version of the software ready soon.

Appendix

Notation for the QT Model

y : ($y_1, y_2, \dots, y_p$) = a set of p random QT phenotype values; note that this means there are p phenotypes. From this point forward, we shall use the term phenotype to mean a continuous random variable, represented by the notation y_i .

n_A : Number of affected individuals;

n_U : Number of unaffected individuals.

Note that we use the term “affected” throughout this work. We could also use the term “case.” We make the same statement for “unaffected” and “control.”

r : Ratio n_U/n_A .

Indices

$1 \leq i \leq p$: Index for phenotype (see above);

$0 \leq k \leq 2$: Index for genotype at the SNP locus; this value is the number of disease or increaser alleles in the SNP genotype.

Genetic Model Parameters

σ_p^2 , $1 \leq i \leq p$: QTL variance of the phenotype y_i , that is, its contribution to the variance of the population’s i -th QT from the QTL. Note that this quantity is the genetic component of the population phenotype variance (specified in this work as $N(0, 1)$).

$\sigma_{R_i}^2$, $1 \leq i \leq p$: Error variance of the phenotype y_i ; using Fisher’s partitioning [104], we have $\sigma_{R_i}^2 = 1 - \sigma_p^2$. Note that the error variance is the common (phenotype-specific) variance for each of the normal components that make up the i -th mixture distribution.

δ_i , $1 \leq i \leq p$: Dominance of the disease allele for the phenotype y_i ; in this work, we restrict δ_i to the range $-1 \leq \delta_i \leq 1$, although in theory the dominance may range between $-\infty$ and ∞ [105].

p_d : Frequency of the disease (“increaser”) allele at the SNP locus of interest;

p_+ : Frequency of the wild-type (“null”) allele at the SNP locus of interest; note that $p_d + p_+ = 1$.

Note that the parameters p_d and p_+ should not be confused with the number of phenotypes p .

a_i , $1 \leq i \leq p$: Additive term for the phenotype y_i ;

$\tau_i = \delta_i/a_i$, $1 \leq i \leq p$: Dominant-additive ratio for the phenotype y_i ;

m_i , $1 \leq i \leq p$: Mean term for the phenotype y_i .

ρ_{ij} : Correlation between the variables y_i and y_j .

w_k , $0 \leq k \leq 2$: Weight of the k -th (coded) genotype in the LTT.

From Fisher’s work [104, 105], we can compute the means μ_{ik} from the dominance δ_i and the disease allele frequency p_d . Fisher shows:

$$\text{I. } a_i = \sqrt{\frac{\sigma_i^2}{2p_d p_+ (1 + \tau_i(p_+ - p_d))^2 + 4(p_d p_+ \tau_i)^2}},$$

$$\text{II. } \delta_i = \tau_i a_i,$$

$$\text{III. } m_i = (-1)[(p_d)^2 a_i + 2p_d p_+ \delta_i - (p_+)^2 a_i],$$

$$\text{IV. } \mu_{i0} = m_i - a_i$$

$$\mu_{i1} = m_i + \delta_i$$

$$\mu_{i2} = m_i + a_i,$$

V. π_k , $0 \leq k \leq 2$: Mixing proportion for the component-distribution $N(\mu_{ik}, \sigma_{R_i}^2)$, determined by the genotype frequencies at the trait locus; because we are studying pleiotropy, the mixing proportions are independent of the phenotype index i . Note that $N(\mu_{ik}, \sigma_{R_i}^2)$ is a univariate normal distribution with the mean μ_{ik} and the variance $\sigma_{R_i}^2$.

Furthermore, as documented by Lynch and Walsh [105] (among others), the genetic variance σ_i^2 may be decomposed into the sum of an additive variance component ($\sigma_{a_i}^2$) and a dominance variance component ($\sigma_{\delta_i}^2$). As Lynch and Walsh report:

$$\text{A. } \sigma_{a_i}^2 = 2p_d p_+ \alpha^2, \text{ where } \alpha = [a_i + \delta_i(p_+ - p_d)];$$

$$\text{B. } \sigma_{\delta_i}^2 = (2p_d p_+ \delta_i)^2.$$

From these equations, it is straightforward to see that the genetic variance for the i -th phenotype is a function of a_i , the additive term for the phenotype y_i , the disease allele frequency p_d , and the dominance δ_i .

Acknowledgements

This study was supported by a grant from the National Institute of Mental Health (R01MH092293 to G.A.H.) and the New Jersey Center for Tourette Syndrome and Associated Disorders (to G.A.H.). The authors gratefully acknowledge the Associate Editor and 2 anonymous reviewers, whose comments substantially improved the quality of our manuscript.

References

- 1 Stearns FW: One hundred years of pleiotropy: a retrospective. *Genetics* 2010;186:767–773.
- 2 Didion JP, de Villena FPM: Deconstructing *Mus gemischus*: advances in understanding ancestry, structure, and variation in the genome of the laboratory mouse. *Mamm Genome* 2013;24:1–20.
- 3 Khalili H, Gong J, Brenner H, Austin TR, Hutter CM, et al: Identification of a common variant with potential pleiotropic effect on risk of inflammatory bowel disease and colorectal cancer. *Carcinogenesis* 2015;36:999–1007.
- 4 Cheng I, Kocarnik JM, Dumitrescu L, Lindor NM, Chang-Claude J, et al: Pleiotropic effects of genetic risk variants for other cancers on colorectal cancer risk: PAGE, GECCO and CCFR consortia. *Gut* 2014;63:800–807.
- 5 Trbojević Akmačić I, Ventham NT, Theodoratou E, Vučković F, Kennedy NA, et al: Inflammatory bowel disease associates with proinflammatory potential of the immunoglobulin G glycome. *Inflamm Bowel Dis* 2015;21:1237–1247.
- 6 Andreassen OA, Desikan RS, Wang Y, Thompson WK, Schork AJ, et al: Abundant genetic overlap between blood lipids and immune-mediated diseases indicates shared molecular genetic mechanisms. *PLoS One* 2015;10:e0123057.
- 7 Chang D, Gao F, Slavney A, Ma L, Waldman YY, et al: Accounting for eXcentricities: analysis of the X chromosome in GWAS reveals X-linked genes implicated in autoimmune diseases. *PLoS One* 2014;9:e113684.
- 8 Li C, Yang C, Gelernter J, Zhao H: Improving genetic risk prediction by leveraging pleiotropy. *Hum Genet* 2014;133:639–650.
- 9 Lauc G, Huffman JE, Pučić M, Zgaga L, Adamczyk B, et al: Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers. *PLoS Genet* 2013;9:e1003225.
- 10 Ramos PS, Criswell LA, Moser KL, Comeau ME, Williams AH, et al: A comprehensive analysis of shared loci between systemic lupus erythematosus (SLE) and sixteen autoimmune diseases reveals limited genetic overlap. *PLoS Genet* 2011;7:e1002406.
- 11 Proitsi P, Lupton MK, Velayudhan L, Newhouse S, Fogh I, et al: Genetic predisposition to increased blood cholesterol and triglyceride lipid levels and risk of Alzheimer disease: a Mendelian randomization analysis. *PLoS Med* 2014;11:e1001713.
- 12 Proitsi P, Lupton MK, Velayudhan L, Hunter G, Newhouse S, et al: Alleles that increase risk for type 2 diabetes mellitus are not associated with increased risk for Alzheimer's disease. *Neurobiol Aging* 2014;35:2883.e3–2883.e10.
- 13 Evans S, Dowell NG, Tabet N, Tofts PS, King SL, Rusted JM: Cognitive and neural signatures of the APOE E4 allele in mid-aged adults. *Neurobiol Aging* 2014;35:1615–1623.
- 14 Adeosun SO, Hou X, Zheng B, Stockmeier C, Ou X, et al: Cognitive deficits and disruption of neurogenesis in a mouse model of apolipoprotein E4 domain interaction. *J Biol Chem* 2014;289:2946–2959.
- 15 Douet V, Chang L, Cloak C, Ernst T: Genetic influences on brain developmental trajectories on neuroimaging studies: from infancy to young adulthood. *Brain Imaging Behav* 2014;8:234–250.
- 16 van Blitterswijk M, Baker MC, DeJesus-Hernandez M, Ghidoni R, Benussi L, et al: C9ORF72 repeat expansions in cases with previously identified pathogenic mutations. *Neurology* 2013;81:1332–1341.
- 17 Buřill E, Blesa R, Augustí J: Alzheimer's disease: an evolutionary approach. *J Anthropol Sci* 2013;91:135–157.
- 18 Jin SC, Pastor P, Cooper B, Cervantes S, Benitez BA, et al: Pooled-DNA sequencing identifies novel causative variants in *PSEN1*, *GRN* and *MAPT* in a clinical early-onset and familial Alzheimer's disease Ibero-American cohort. *Alzheimers Res Ther* 2012;4:34.
- 19 Albin RL: Antagonistic pleiotropy, mutation accumulation, and human genetic disease. *Genetica* 1993;91:279–286.
- 20 Sun QB, Zhang KZ, Cheng TO, Li SL, Lu BX, et al: Marfan syndrome in China: a collective review of 564 cases among 98 families. *Am Heart J* 1990;120:934–948.
- 21 Pyeritz RE: Pleiotropy revisited: molecular explanations of a classic concept. *Am J Med Genet* 1989;34:124–134.
- 22 Baumgartner C, Mátyás G, Steinmann B, Eberle M, Stein JL, Baumgartner D: A bioinformatics framework for genotype-phenotype correlation in humans with Marfan syndrome caused by *FBN1* gene mutations. *J Biomed Inform* 2006;39:171–183.
- 23 Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW: Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet* 2013;14:483–495.
- 24 Mitra SK: On the limiting power function of the frequency chi-square test. *Ann Math Stat* 1958;29:1221–1233.
- 25 Slager SL, Schaid DJ: Case-control studies of genetic markers: power and sample size approximations for Armitage's test for trend. *Hum Hered* 2001;52:149–153.
- 26 Chapman DG, Nam JM: Asymptotic power of chi square tests for linear trends in proportions. *Biometrics* 1968;24:315–327.
- 27 Freidlin B, Zheng G, Li Z, Gastwirth JL: Trend tests for case-control studies of genetic markers: power, sample size and robustness. *Hum Hered* 2002;53:146–152.
- 28 Menashe I, Rosenberg PS, Chen BE: PGA: power calculator for case-control genetic association analyses. *BMC Genet* 2008;9:36.
- 29 Barrenäs F, Chavali S, Alves AC, Coin L, Jarvelin MR, et al: Highly interconnected genes in disease-specific networks are enriched for disease-associated polymorphisms. *Genome Biol* 2012;13:R46.
- 30 Chung D, Yang C, Li C, Gelernter J, Zhao H: GPA: a statistical approach to prioritizing GWAS results by integrating pleiotropy and annotation. *PLoS Genet* 2014;10:e1004787.
- 31 Darabos C, Harmon SH, Moore JH: Using the bipartite human phenotype network to reveal pleiotropy and epistasis beyond the gene. *Pac Symp Biocomput* 2014:188–199.
- 32 Darabos C, Moore JH: Genome-wide epistasis and pleiotropy characterized by the bipartite human phenotype network. *Methods Mol Biol* 2015;1253:269–283.
- 33 Hartley SW, Sebastiani P: PleioGRiP: genetic risk prediction with pleiotropy. *Bioinformatics* 2013;29:1086–1088.
- 34 He Q, Avery CL, Lin DY: A general framework for association tests with multivariate traits in large-scale genomics studies. *Genet Epidemiol* 2013;37:759–767.
- 35 Huang J, Johnson AD, O'Donnell CJ: PRIMe: a method for characterization and evaluation of pleiotropic regions from multiple genome-wide association studies. *Bioinformatics* 2011;27:1201–1206.
- 36 Lee SH, Yang J, Goddard ME, Visscher PM, Wray NR: Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* 2012;28:2540–2542.
- 37 Li Q, Hu J, Ding J, Zheng G: Fisher's method of combining dependent statistics using generalizations of the gamma distribution with applications to genetic pleiotropic associations. *Biostatistics* 2014;15:284–295.
- 38 Liley J, Wallace C: A pleiotropy-informed Bayesian false discovery rate adapted to a shared control design finds new disease associations from GWAS summary statistics. *PLoS Genet* 2015;11:e1004926.
- 39 Matise TC, Ambite JL, Buyske S, Carlson CS, Cole SA, et al: The next PAGE in understanding complex traits: design for the analysis of Population Architecture Using Genetics and Epidemiology (PAGE) Study. *Am J Epidemiol* 2011;174:849–859.
- 40 Park SH, Lee JY, Kim S: A methodology for multivariate phenotype-based genome-wide association studies to mine pleiotropic genes. *BMC Syst Biol* 2011;5(suppl 2):S13.
- 41 Seoane JA, Campbell C, Day IN, Casas JP, Gaunt TR: Canonical correlation analysis for gene-based pleiotropy discovery. *PLoS Comput Biol* 2014;10:e1003876.
- 42 Sivakumaran S, Agakov F, Theodoratou E, Prendergast JG, Zgaga L, et al: Abundant pleiotropy in human complex diseases and traits. *Am J Hum Genet* 2011;89:607–618.

- 43 Wu B, Pankow JS: Statistical methods for association tests of multiple continuous traits in genome-wide association studies. *Ann Hum Genet* 2015;79:282–293.
- 44 Yan T, Li Q, Li Y, Li Z, Zheng G: Genetic association with multiple traits in the presence of population stratification. *Genet Epidemiol* 2013;37:571–580.
- 45 Zhang Q, Feitosa M, Borecki IB: Estimating and testing pleiotropy of single genetic variant for two quantitative traits. *Genet Epidemiol* 2014;38:523–530.
- 46 Pendergrass SA, Verma A, Okula A, Hall MA, Crawford DC, Ritchie MD: Phenome-wide association studies: embracing complexity for discovery. *Hum Hered* 2015;79:111–123.
- 47 Schifano ED, Li L, Christiani DC, Lin X: Genome-wide association analysis for multiple continuous secondary phenotypes. *Am J Hum Genet* 2013;92:744–759.
- 48 Peterson CB, Bogomolov M, Benjamini Y, Sabbati C: Many phenotypes without many false discoveries: error controlling strategies for multitrait association studies. *Genet Epidemiol* 2016;40:45–56.
- 49 Ray D, Pankow JS, Basu S: USAT: a unified score-based association test for multiple phenotype-genotype analysis. *Genet Epidemiol* 2016;40:20–34.
- 50 Vsevolozhskaya OA, Zaykin DV, Barondess DA, Tong X, Jadhav S, Lu Q: Uncovering local trends in genetic effects of multiple phenotypes via functional linear models. *Genet Epidemiol* 2016;40:210–221.
- 51 Majumdar A, Haldar T, Witte JS: Determining which phenotypes underlie a pleiotropic signal. *Genet Epidemiol* 2016;40:366–381.
- 52 Baurecht H, Hotze M, Rodríguez E, Manz J, Weidinger S, et al: Compare and Contrast Meta-Analysis (CCMA): a method for identification of pleiotropic loci in genome-wide association studies. *PLoS One* 2016;11:e0154872.
- 53 Bowden J, Davey Smith G, Haycock PC, Burgess S: Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol* 2016;40:304–314.
- 54 Denny JC, Bastarache L, Roden DM: Phenome-wide association studies as a tool to advance precision medicine. *Annu Rev Genomics Hum Genet* 2016;17:353–373.
- 55 Hall MA, Moore JH, Ritchie MD: Embracing complex associations in common traits: critical considerations for precision medicine. *Trends Genet* 2016;32:470–484.
- 56 Liang X, Wang Z, Sha Q, Zhang S: An adaptive Fisher's combination method for joint analysis of multiple phenotypes in association studies. *Sci Rep* 2016;6:34323.
- 57 Park H, Li X, Song YE, He KY, Zhu X: Multivariate analysis of anthropometric traits using summary statistics of genome-wide association studies from GIANT consortium. *PLoS One* 2016;11:e0163912.
- 58 Verma A, Leader JB, Verma SS, Frase A, Wallace J, et al: Integrating clinical laboratory measures and ICD-9 code diagnoses in phenome-wide association studies. *Pac Symp Biocomput* 2016;21:168–179.
- 59 Wang X, Byars SG, Stearns SC: Genetic links between post-reproductive lifespan and family size in Framingham. *Evol Med Public Health* 2013;2013:241–253.
- 60 Knowles EE, McKay DR, Kent JW Jr, Sprooten E, Carless MA, et al: Pleiotropic locus for emotion recognition and amygdala volume identified using univariate and bivariate linkage. *Am J Psychiatry* 2015;172:190–199.
- 61 Schifano ED, Li L, Christiani DC, Lin X: Genome-wide association analysis for multiple continuous secondary phenotypes. *Am J Hum Genet* 2013;92:744–759.
- 62 Curran JE, McKay DR, Winkler AM, Olvera RL, Carless MA, et al: Identification of pleiotropic genetic effects on obesity and brain anatomy. *Hum Hered* 2013;75:136–143.
- 63 Hokanson JE, Langefeld CD, Mitchell BD, Lange LA, Goff DC Jr, et al: Pleiotropy and heterogeneity in the expression of atherogenic lipoproteins: the IRAS Family Study. *Hum Hered* 2003;55:46–50.
- 64 Miscimarra L, Stein C, Millard C, Kluge A, Cartier K, et al: Further evidence of pleiotropy influencing speech and language: analysis of the DYX8 region. *Hum Hered* 2007;63:47–58.
- 65 Morton NE, Lalouel JM: Resolution of linkage for irregular phenotype systems. *Hum Hered* 1981;31:3–7.
- 66 Njajou OT, Alizadeh BZ, Aulchenko Y, Zillickens MC, Pols HA, et al: Heritability of serum iron, ferritin and transferrin saturation in a genetically isolated population, the Erasmus Rucphen Family (ERF) Study. *Hum Hered* 2006;61:222–228.
- 67 Li Z, Möttönen J, Sillanpää MJ: A robust multiple-locus method for quantitative trait locus analysis of non-normally distributed multiple traits. *Heredity (Edinb)* 2015;115:556–564.
- 68 Lee D, Williamson VS, Bigdeli TB, Riley BP, Fanous AH, et al: JEPeg: a summary statistics based tool for gene-level joint testing of functional variants. *Bioinformatics* 2015;31:1176–1182.
- 69 Yuan Z, Zhang X, Li F, Zhao J, Xue F: Comparing partial least square approaches in a gene- or region-based association study for multiple quantitative phenotypes. *Hum Biol* 2014;86:51–58.
- 70 Fu G, Saunders G, Stevens J: Holm multiple correction for large-scale gene-shape association mapping. *BMC Genet* 2014;15(suppl 1):S5.
- 71 Yoo YJ, Sun L, Bull SB: Gene-based multiple regression association testing for combined examination of common and low frequency variants in quantitative trait analysis. *Front Genet* 2013;4:233.
- 72 Ma L, Clark AG, Keinan A: Gene-based testing of interactions in association studies of quantitative traits. *PLoS Genet* 2013;9:e1003321.
- 73 Fan R, Lo SH: A robust model-free approach for rare variants association studies incorporating gene-gene and gene-environmental interactions. *PLoS One* 2013;8:e83057.
- 74 Clarke GM, Rivas MA, Morris AP: A flexible approach for the analysis of rare variants allowing for a mixture of effects on binary or quantitative traits. *PLoS Genet* 2013;9:e1003694.
- 75 Zhang F, Guo X, Wu S, Han J, Liu Y, et al: Genome-wide pathway association studies of multiple correlated quantitative phenotypes using principle component analyses. *PLoS One* 2012;7:e53320.
- 76 Korte A, Vilhjálmsson BJ, Segura V, Platt A, Long Q, Nordborg M: A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nat Genet* 2012;44:1066–1071.
- 77 Li M, Ye C, Fu W, Elston RC, Lu Q: Detecting genetic interactions for quantitative traits with U-statistics. *Genet Epidemiol* 2011;35:457–468.
- 78 Yang F, Tang Z, Deng H: Bivariate association analysis for quantitative traits using generalized estimation equation. *J Genet Genomics* 2009;36:733–743.
- 79 Kent JW Jr: Analysis of multiple phenotypes. *Genet Epidemiol* 2009;33(suppl 1):S33–S39.
- 80 Hu Y, Jason S, Wang Q, Pan Y, Zhang X, et al: Regression-based approach for testing the association between multi-region haplotype configuration and complex trait. *BMC Genet* 2009;10:56.
- 81 Fang M, Liu S, Jiang D: Bayesian composite model space approach for mapping quantitative trait loci in variance component model. *Behav Genet* 2009;39:337–346.
- 82 Wei Z, Li M, Rebbeck T, Li H: U-statistics-based tests for multiple genes in genetic association studies. *Ann Hum Genet* 2008;72:821–833.
- 83 Servin B, Stephens M: Imputation-based analysis of association studies: candidate regions and quantitative traits. *PLoS Genet* 2007;3:e114.
- 84 Fan R, Jung J, Jin L: High-resolution association mapping of quantitative trait loci: a population-based approach. *Genetics* 2006;172:663–686.
- 85 Lange C, DeMeo DL, Laird NM: Power and design considerations for a general class of family-based association tests: quantitative traits. *Am J Hum Genet* 2002;71:1330–1341.
- 86 Tyler AL, McGarr TC, Beyer BJ, Frankel WN, Carter GW: A genetic interaction network model of a complex neurological disease. *Genes Brain Behav* 2014;13:831–840.
- 87 Purcell S, Cherny SS, Sham PC: Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003;19:149–150.
- 88 Gordon D, Haynes C, Blumenfeld J, Finch SJ: PAWE-3D: visualizing power for association with error in case-control genetic studies of complex traits. *Bioinformatics* 2005;21:3935–3937.

- 89 The Marfan Foundation. 2016. <http://www.marfan.org/dx/rules>.
- 90 Loeys BL, Dietz HC, Braverman AC, Callewaert BL, De Backer J, et al: The revised Ghent nosology for the Marfan syndrome. *J Med Genet* 2010;47:476–485.
- 91 American Psychiatric Association: DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders, ed 4, text rev. Washington, American Psychiatric Association, 2000.
- 92 Boghosian-Sell L, Comings DE, Overhauser J: Tourette syndrome in a pedigree with a 7;18 translocation: identification of a YAC spanning the translocation breakpoint at 18q22.3. *Am J Hum Genet* 1996;59:999–1005.
- 93 Díaz-Anzaldúa A, Rivière JB, Dubé MP, Joober R, Saint-Onge J, et al: Chromosome 11-q24 region in Tourette syndrome: association and linkage disequilibrium study in the French Canadian population. *Am J Med Genet A* 2005;138A:225–228.
- 94 Saïdou AA, Thuillet AC, Couderc M, Mariac C, Vigouroux Y: Association studies including genotype by environment interactions: prospects and limits. *BMC Genet* 2014;15:3.
- 95 Xiao J, Wang X, Hu Z, Tang Z, Xu C: Multivariate segregation analysis for quantitative traits in line crosses. *Heredity (Edinb)* 2007;98:427–435.
- 96 Liu J, Liu Y, Liu X, Deng HW: Bayesian mapping of quantitative trait loci for multiple complex traits with the use of variance components. *Am J Hum Genet* 2007;81:304–320.
- 97 Kraft P, de Andrade M: Group 6: pleiotropy and multivariate analysis. *Genet Epidemiol* 2003;25(suppl 1):S50–S56.
- 98 Bensen JT, Lange LA, Langefeld CD, Chang BL, Bleecker ER, et al: Exploring pleiotropy using principal components. *BMC Genet* 2003;4(suppl 1):S53.
- 99 Lebreton CM, Visscher PM, Haley CS, Semikhodskii A, Quarrie SA: A nonparametric bootstrap method for testing close linkage vs pleiotropy of coincident quantitative trait loci. *Genetics* 1998;150:931–943.
- 100 Almasy L, Dyer TD, Blangero J: Bivariate quantitative trait linkage analysis: pleiotropy versus co-incident linkages. *Genet Epidemiol* 1997;14:953–958.
- 101 Jiang C, Zeng ZB: Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* 1995;140:1111–1127.
- 102 Warne RT: A primer on multivariate analysis of variance (MANOVA) for behavioral scientists. *Pract Assess Res Eval* 2014;19:1–10.
- 103 Olson CL: On choosing a test statistic in multivariate analysis of variance. *Psychol Bull* 1976;83:579–586.
- 104 Fisher RA: The correlation between relatives on the supposition of Mendelian inheritance. *Trans R Soc Edinb* 1918;52:399–433.
- 105 Lynch M, Walsh B: *Genetics and Analysis of Quantitative Traits*. Sunderland, Sinauer, 1998.
- 106 R Development Core Team: *R: A Language and Environment for Statistical Computing*. Vienna, R Foundation for Statistical Computing, 2012.
- 107 O'Brien RG, Shieh G: Pragmatic, unifying algorithm gives power probabilities for common *F* tests of the multivariate general linear hypothesis (technical report). 1999. <http://www.bio.ri.ccf.org/UnifyPow>.
- 108 Box GEP, Hunter GS, Hunter WG: *Statistics for Experimenters: Design, Discovery, and Innovation*, ed 2. Hoboken, Wiley & Sons, 2005.
- 109 Lehmann EL, Romano JP: *Testing Statistical Hypotheses*, ed 3. New York, Springer, 2010.