Genetics: Early Online, published on April 13, 2016 as 10.1534/genetics.115.186502

Powerful and adaptive testing for multi-trait and multi-SNP associations with GWAS and sequencing data

Junghi Kim[1], Yiwei Zhang[1], Wei Pan[1],

for the Alzheimer's Disease Neuroimaging Initiative[2]

[1]*Division of Biostatistics, University of Minnesota, Minneapolis, MN 55455, USA*

December 22, 2015; revised March 11, 2016

Running title: Testing for multi-trait and multi-SNP associations

Correspondence author: Wei Pan

Telephone: (612) 626-2705

Fax: (612) 626-0660

Email: weip@biostat.umn.edu

Address: Division of Biostatistics, MMC 303,

School of Public Health, University of Minnesota,

Minneapolis, Minnesota 55455–0392, U.S.A.

[2] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (`adni.loni.usc.edu`). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: `http://adni.loni.usc.edu/wp-content/uploads/howtoapply/ADNIAcknowledgementList.pdf`.

Abstract

Testing for genetic association with multiple traits has become increasingly important, not only because of its potential to boost statistical power, but also for its direct relevance to applications. For example, there is accumulating evidence showing that some complex neurodegenerative and psychiatric diseases like Alzheimer's are due to disrupted brain networks, for which it would be natural to identify genetic variants associated with a disrupted brain network, represented as a set of multiple traits, one for each of multiple brain regions of interest (ROIs). In spite of its promise, testing for multivariate trait associations is challenging: if not appropriately used, its power can be much lower than testing on each univariate trait separately (with a proper control for multiple testing). Furthermore, differing from most existing methods for single SNP-multiple trait associations, we consider SNP set-based association testing to decipher complicated joint effects of multiple SNPs on multiple traits. Because the power of a test critically depends on several unknown factors such as the proportions of associated SNPs and of traits, we propose a highly adaptive test at both the SNP and trait levels, giving higher weights to those likely associated SNPs and traits, to yield high power across a wide spectrum of situations. We illuminate on relationships among the proposed and some existing tests, showing that the proposed test covers several existing tests as special cases. We compare the performance of the new test with several existing tests using both simulated and real data. The methods were applied to structural MRI data drawn from Alzheimer's Disease Neuroimaging Initiative (ADNI) to identify genes associated with grey matter atrophy in the human brain default mode network (DMN). For GWAS, genes $AMOTL1$ on chromosome 11 and $APOE$ on chromosome 19 were discovered by the new test to be significantly associated with DMN. Notably, gene $AMOTL1$ was not detected by single SNP-based analyses. To our knowledge, $AMOTL1$ has not been highlighted in other AD studies before, though it was indicated to be related to cognitive impairment. The proposed method is also applicable to rare variants in sequencing data and can be extended to pathway analysis.

Keywords: adaptive association test; ADNI; default mode network; gene-based test; imaging genetics; multiple traits

Introduction

Alzheimer's disease (AD) (MIM 104300) is the most common neurodegenerative disease, and every 67 seconds, someone in U.S develops AD (Alzheimer's Association 2015a). Currently there is no cure for AD, and most cases are diagnosed in the late stage of the disease. It is projected that the number of Americans of age 65 and older with AD will increase from 5.1 million in 2015 to 13.5 million in 2050, an growth from an estimated 11% of the US senior population in 2015 to 16% in 2050, costing over $1.1 trillion in 2050 (Alzheimer's Association 2015b). To advance our understanding of the initiation, progression and etiology of AD, Alzheimer's Disease Neuroimaging Initiative (ADNI) was started in 2004 and is being continued since, collecting extensive clinical, genomic and multi-modal imaging data (Shen et al. 2014). Many other genetic studies have been conducted, identifying multiple common and rare variants, shedding light on pathogenic mechanisms of AD (Marei et al. 2015; Saykin et al. 2015). In particular, the APOEε4 allele has been consistently shown to be associated with AD. However, only 50% of AD patients carry an APOEε4 allele, suggesting the existence of other genetic variants contributing to risk for the disease (Karch et al. 2014). A recent study indicates that 33% of total AD phenotypic variance is explained by common variants; APOE alone explains 6% and other known markers 2%, meaning more than 25% of phenotypic variance remains unexplained by known common variants (Ridge et al. 2013). Hence, as for other common and complex diseases and traits, many more genetic factors underlying late onset AD are waiting to be discovered. One obvious but costly approach is to have a larger sample size. Alternatively, more powerful analysis methods are urgently needed. For example, in contrast to the popular single SNP-based analysis, novel gene- and pathway-based analyses may be more powerful in discovering additional causal variants. As demonstrated by Jones et al. (2010), jointly analyzing functionally related SNPs sheds new light on the relatedness of immune regulation, energy metabolism and protein degradation to the etiology of AD. The reason is due to the well-known genetic heterogeneity and small effect sizes of individual common variants, as observed from published GWAS results (Manolio et al. 2009). To boost power in identifying aggregate effects of multiple SNPs, it may be promising to conduct association analysis at the SNP-set (or gene) level, rather than at the individual SNP level.

Another strategy is to use multiple endophenotypes, intermediate between genetics and the

disease, for their potential to have stronger associations with genetic variants. In addition to boosting power, the use of intermediate phenotypes may provide important clues about causal pathways to the disease (Schifano et al. 2013; Maity et al. 2012). A recent GWAS demonstrated the effectiveness of the strategy: some risk genes such as $FRMD6$, were first identified to be associated with some neuroimaging intermediate phenotypes (e.g. hippocampal atrophy) (Shen et al. 2014), then were later validated to be associated with AD (Hong et al. 2012; Sherva et al. 2014). A possibly useful but under-utilized intermediate phenotype is the brain default mode network (DMN), consisting of several brain regions of interest (ROIs) remaining active in the resting state. Brain activity in DMN may explain the etiology of AD (Metin et al. 2015), and is a plausible indicator for incipient AD (Damoiseaux et al 2013; Greicius et al. 2004; He et al. 2009; Jones et al. 2011; Balthazar et al. 2014). Since there is growing evidence that genetic factors play a role in aberrant default mode connectivity (Glahn et al. 2009), it may be substantially more powerful to detect genetic variants associated with DMN, a set of multiple intermediate phenotypes, than with AD.

Here we discuss gene-based multi-trait analysis, aiming at discovering genes associated with multiple traits such as DMN. To date, several but not many methods have been proposed for gene-based multi-trait analysis (Guo et al. 2013; Van der Sluis et al. 2015; Maity et al. 2014; Wang et al. 2015). The simplest way is to use the minimum p-value (minP) test based on the most significant single SNP–single trait association, which however may lose power in the presence of multiple weak associations between multiple SNPs and multiple traits. Some methods, such as Van der Sluis et al. (2015) and M-TopQ25Stat (Guo et al. 2013), only utilize a few top association signals among the pairwise single SNP-single trait associations. Some methods based on principal components analysis (PCA) or principal components of heritability (PCH), originally proposed for multiple SNPs and a single trait (Wang and Abbott 2007; Klei et al. 2008), may be also applied. However, these methods and canonical correlation analysis (CCA) (Tang and Ferreira, 2012) make use of only one or few top components, thus they share the same weakness of power loss in the presence of multiple associations; furthermore, the number of PCs may be difficult to determine (Aschard et al. 2014; Huang et al. 2014). Another extreme is the burden test (Shen et al. 2010; Guo et al. 2013; Mukherjee et al. 2014), which is powerful in the presence of a dense association pattern, in which most SNP-trait pairs are associated with almost equal effect sizes and directions;

otherwise, e.g. when the association directions of some SNP-trait pairs are different, it does not perform well (as well known for analysis of rare variants). A compromise between the above two extremes is a variance-component test (Maity et al. 2012; Wang et al. 2013), which is more robust to association density/sparsity and varying association directions. Nevertheless, as shown in the context for multiple rare variants and a single trait (Pan et al 2014), it may still suffer from power loss in the presence of more sparse association patterns (i.e. when there are a fewer associated SNP-trait pairs). A fundamental challenge in multivariate analysis is the lack of a uniformly most powerful test: any test may be powerful in some situations, but not in others. Nevertheless, we aim to construct an adaptive test such that it can maintain high power, not necessarily highest power, across a wide range of scenarios. In particular, the proposed test is adaptive at both the SNP and trait levels. Its key feature is the use of a weighting scheme to yield robust statistical power no matter whether the true and unknown association pattern is dense or sparse (or in whatever directions), and the weight is determined data-adaptively. In addition, some chosen weights correspond to several existing tests, including a burden test and a variance-component test. Therefore, the high power range of the proposed test covers those of the burden test and the variance-component test. Moreover, the proposed test is based on the general framework of the generalized estimating equations (GEE), hence it is flexible with the capability to incorporate covariates and various types of traits (Liang and Zeger, 1986). It also avoids a difficulty in correctly specifying a joint multivariate distribution or likelihood for a set of multiple traits. Furthermore, we extend the proposed method to pathway analysis, in which it is adaptive to possibly varying gene-level associations.

We will compare the performance of the new test with several existing tests using both simulated and real data. The methods were applied to structural MRI data drawn from the ADNI to identify genes associated with DMN. In the GWAS, 277,527 SNPs were mapped to 17,557 genes, among which genes $AMOTL1$ on chromosome 11 and $APOE$ on chromosome 19 were discovered by the new test to be significantly associated with DMN. Notably, gene $AMOTL1$ was not detected by single SNP-based analyses. We also illustrate the application of the methods to the ADNI whole-genome sequencing (WGS) data, though none significant genes were identified, presumably due to a relatively small sample size.

In the following, we briefly review GEE and an existing method before introducing the new test

in Materials and Methods. In Results, the new and several existing methods are compared with applications to the ADNI data and simulated data mimicking the ADNI data. We end with a short summary of the conclusions.

Materials and Methods

Review

Generalized estimating equations

Suppose for each individual $i = 1, ..., n$, we observe k traits $Y_i = (y_{i1}, ..., y_{ik})'$, q covariates $z_i = (z_{i1}, ..., z_{iq})'$ and a set of single nucleotide polymorphisms (SNPs) $x_i = (x_{i1}, ..., x_{ip})'$, with $x_{ij} \in \{0, 1, 2\}$. Denote $X_i = I \otimes x_i^{'}$ and $Z_i = I \otimes (1, z_i^{'})$, where I is a $k \times k$ identity matrix, and $\otimes$ represents the Kronecker product. We model the mean of the phenotypes $E(Y_i|X_i, Z_i) = \mu_i$, using a marginal generalized linear model

$$g(\mu_i) = Z_i\varphi + X_i\beta = H_i\theta \tag{1}$$

with $H_i = \begin{pmatrix} Z_i & X_i \end{pmatrix}$, parameters $\theta = (\varphi^{'}, \beta^{'})'$, and a link function $g(.)$. The regression coefficients $\beta = (\beta_{11}, ..., \beta_{p1}, ..., \beta_{1k}, ..., \beta_{pk})'$ is a $pk \times 1$ vector, in which β_{jt} represents the effect of the jth SNP on the tth trait, while the element φ_{st} of $\varphi = (\varphi_{11}, ..., \varphi_{(q+1)1}, ..., \varphi_{1k}, ..., \varphi_{(q+1)k})'$ is the effect size of the sth covariate on the tth trait. Liang and Zeger (1986) proposed estimating φ and β by solving the generalized estimating equations (GEE):

$$U_\theta = \sum_{i=1}^{n} D_i^{'} V_i^{-1} (Y_i - \mu_i) = 0 \tag{2}$$

with $D_i = \partial\mu_i/\partial\theta'$ and $V_i = \phi A_i^{1/2} R_w(\alpha) A_i^{1/2}$, where ϕ is a dispersion parameter, $A_i = \text{diag}\{v(\mu_{i1}), ..., v(\mu_{ik})\}$ models the variances with a variance function $v(\mu_i)$, and $R_w(\alpha)$ is a working correlation matrix with possibly some unknown parameters α. Specifically, for quantitative traits (Y_i) with the identity link function (or more generally, for any generalized linear model with a canonical link function),

the score vector U_θ and its variance-covariance matrix $Cov(U_\theta)$ are

$$U_\theta = (U_\varphi^{'}, U_\beta^{'})^{'} = \sum_{i=1}^{n} (Z_i \;\; X_i)^{'} R_w^{-1} (Y_i - \mu_i),$$
$$Cov(U_\theta) = \sum_{i=1}^{n} (Z_i \;\; X_i)^{'} R_w^{-1} (Y_i - \mu_i)(Y_i - \mu_i)' R_w^{-1} (Z_i \;\; X_i).$$

The covariance matrix can be partitioned according to the score components for φ and β: $Cov(U_\theta) = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}$. For convenience, the working independence model is often used with R_w being as an identity matrix $I_{k\times k}$, as done in this paper unless specified otherwise.

Our primary concern is to test for overall genetic effects with H_0: $\beta = 0$, while treating φ as nuisance parameters. To perform the score test, we evaluate the equation (1) under H_0. Under H_0, we have $g(\mu_i) = Z_i\varphi$, and the estimate of φ, denoted as $\widehat{\varphi}$, is the solution to the generalized score equation $U_{\varphi,\beta=0} = \sum_{i=1}^{n} Z_i^{'}(Y_i - \mu_i) = 0$. The marginal mean is estimated by $\widehat{\mu}_i = g(Z_i\widehat{\varphi})^{-1}$.

For testing SNP-set effects, one considers the sub-components of the score vector for β:

$$U_\beta = \sum_{i=1}^{n} X_i^{'}(Y_i - \widehat{\mu}_i). \tag{3}$$

U_β asymptotically follows a multivariate normal distribution $\mathcal{MN}(0, \tilde{\Sigma}_\beta)$ under H_0, where $\tilde{\Sigma}_\beta = V_{22} - V_{21}V_{11}^{-1}V_{12}$. U_β can be written as $U_\beta = (U_{11}, ..., U_{p1}, ..., U_{1k}, ..., U_{pk})'$. Each element U_{jt} measures the association strength between SNP j and trait k for $j = 1, ..., p$ and $t = 1, ..., k$, and is asymptotically proportional to β_{jt} in equation (1). $\beta_{jt} = 0$ implies there is no association between SNP j and trait k; similarly $U_{jt} = 0$ (or small) indicates no (or weak) association between SNP j and trait k.

For testing H_0, the GEE-Score test statistic is defined by

$$\text{GEE-Score} = U_\beta^{'}\tilde{\Sigma}_\beta^{-1}U_\beta.$$

Under H_0, the GEE-Score statistic asymptotically follows a central chi-squared distribution with pk degrees of freedom. When pk is large, this standard score test loses power for large degrees of freedom. Another way to draw inference, especially convenient when combining the score test with other tests as to be discussed later, is to simulate $U_\beta^{(b)} \sim \mathcal{MN}(0, \tilde{\Sigma}_\beta)$ for $b = 1, ..., B$ and

obtain the null statistics GEE-Score$^{(b)} = U_\beta^{(b)\prime}\tilde{\Sigma}_\beta^{-1}U_\beta^{(b)}$. The p-value can be calculated as $P_{\text{Score}} = \sum_{b=1}^{B} I(\text{GEE-Score} \leq \text{GEE-Score}^{(b)})/(B+1)$, where $I(\cdot)$ denotes the indicator function.

For ease of notation, we suppress β and take $U = U_\beta$ and $V = \tilde{\Sigma}_\beta$ hereafter.

An adaptive association test for a single SNP

Zhang et al. (2014) proposed a class of sum of powered score (SPU) tests for testing association between an individual SNP and multiple traits, along with its data-adaptive version (aSPU). The SPU tests are a family of association tests based on the (generalized) score vector in the GEE framework, aiming for at least one of them to be powerful in any given situation. With only a single SNP j, then the score vector reduces to $U = (U_{j1}, ..., U_{jk})'$. The association between the SNP and k traits can be quantified with a test statistic

$$\text{SPU}(\gamma) = \sum_{t=1}^{k} (U_{jt})^\gamma$$

where a candidate integer $\gamma \geq 1$ is to be chosen from a pre-selected parameter set Γ; e.g. $\Gamma = \{1, 2, ..., 8, \infty\}$. The statistical power of an SPU(γ) test depends on the choice of $\gamma \in \Gamma$. When γ is an odd integer, the SPU(γ) test sums up the association signals across all the traits, retaining high power if all or most of the multiple traits have an almost equal effect size in the same association direction. A special case is $\gamma = 1$, giving a burden test commonly used for rare variants. With an even γ, the SPU(γ) test will be more powerful when some traits have different association directions. In particular, the SPU(2) test is the same as the sum of squared score (SSU) test (Pan 2011), closely related to MDMR (McArdle and Anderson 2001), kernel machine regression (KMR) (Liu et al. 2007) and variance-component tests (Tzeng et al. 2011). Furthermore, as γ increases, the SPU test upweights the more strongly associated traits, while reducing the weights on other ones. In particular, when $\gamma \to \infty$ (as an even integer), only the maximum component of the score vector is used and the test statistic is defined as $\text{SPU}(\infty) = \max_{t=1}^{k} |U_{jt}|$. The SPU($\infty$) test is similar to the UminP test (when the variances of the score components are almost equal). To compute the significance of an SPU test, Monte Carlo (MC) simulations (or alternatively, permutations) are used; for $b = 1, ..., B$, the null score $U^{(b)} = (U_{j1}^{(b)}, ..., U_{jk}^{(b)})'$ is generated from $\mathcal{MN}(0, V)$, from which the null statistics $\text{SPU}(\gamma)^{(b)} = \sum_{t=1}^{k} (U_{jt}^{(b)})^\gamma$ can be obtained for each γ. Then the p-value can be

calculated as $p_\gamma = [\sum_{b=1}^{B} I(\text{SPU}(\gamma) \leq \text{SPU}(\gamma)^{(b)}) + 1]/(B+1)$.

However, it is not clear how to choose an optimal γ a priori for given data. Hence, Zhang et al. (2014) proposed an adaptive SPU (aSPU) test to extract association evidence from multiple SPU(γ) tests. The statistic of the aSPU test is the minimum p-value of SPU(γ)'s for some candidate values of γ:

$$\text{aSPU} = \min_{\gamma \in \Gamma} p_\gamma,$$

where p_γ is p-value of SPU(γ). By MC simulations (or permutations), the p-value of aSPU, along with those of all SPU(γ) tests, can be efficiently calculated based on the same set of the null statistics in a single layer.

Existing gene-based tests

We will compare the proposed test with several existing gene-based tests for multiple traits, including multivariate analysis of variance (MANOVA), multivariate distance matrix regression (MDMR) with the Euclidean distance (McArdle and Anderson 2001), multivariate kernel machine regression (KMR) under linear kernel (Maity et al. 2012) and a multivariate functional linear model (MFLM) (Wang et al. 2015). We would note that KMR can be derived based on a ramdom-effects model while MFLM is built on a fixed effect model. For implementation, R package `vegan` was used for MDMR; R code for KMR and MFLM was downloaded from the authors' websites, `http://www4.stat.ncsu.edu/~maity/software.html` and `https://www.nichd.nih.gov/about/org/diphr/bbb/software/fan/Pages/default.aspx` respectively. Since KMR (Maity et al. 2012) was computationally slow, and excluded from the simulation studies.

New Methods

An adaptive test

We introduce a novel gene-based adaptive sum of powered score test for a set of multiple traits, denoted as *aSPUset*, by extending the single SNP-based test of Zhang et al. (2014). Suppose that there are p SNPs in a gene and k traits of interests. Recall that $U = (U_{11}, ..., U_{p1}, ..., U_{1k}, ..., U_{pk})'$ is the generalized score vector of length pk in GEE, and V is the $pk \times pk$ covariance matrix of the score vector; each element of the score, U_{jt} quantifies the association between SNP j and trait t.

In practice, the true and unknown association patterns across multiple SNPs and multiple traits are complex: some SNPs may be associated with some traits, but not with other traits; different SNPs may be associated with different subsets of the traits with varying association strengths and directions. Since the use of non-associated SNPs and traits in a test statistic could reduce the power of the test, we may want to give higher weights to more likely associated SNPs and traits. However, how much to optimally overweight these likely associated SNPs and traits depends on the true association pattern, which is unknown. The aSPUset test employs two positive integer parameters, γ_1 and γ_2, to control the degrees of weighting over the SNPs and over the traits respectively, and the two parameters are chosen data-adaptively. A larger γ_1 puts more weights on the SNPs more likely to be associated with a given trait, while a larger γ_2 upweights the traits more strongly associated with the SNPs.

We build the test statistic as follows. For each trait t, $S(\gamma_1;\ t)$ quantifies the association between the single trait and multiple SNPs, then $\mathrm{SPU}(\gamma_1, \gamma_2)$ combines the single trait-based statistics:

$$S(\gamma_1;\ t) = \big(\sum_{j=1}^{p}(U_{jt})^{\gamma_1}\big)^{1/\gamma_1}, \qquad \mathrm{SPU}(\gamma_1, \gamma_2) = \sum_{t=1}^{k}\big(S(\gamma_1;\ t)\big)^{\gamma_2}. \tag{4}$$

Here candidate integers $\gamma_1 \geq 1$ and $\gamma_2 \geq 1$ are to be chosen from two pre-selected parameter sets Γ_1 and Γ_2. We used $\Gamma_1 = \Gamma_2 = \{1, 2, ..., 8, \infty\}$, due to the good performance in our numerical studies.

In $S(\gamma_1;\ t)$, $(U_{jt})^{\gamma_1}$ can be re-written by an alternative form $(U_{jt})^{\gamma_1} = U_{jt}^{\gamma_1 - 1}U_{jt} = w_{jt}U_{jt}$. $w_{jt} = U_{jt}^{\gamma_1 - 1}$ is a weight for each score element, which reflects the association strength (and direction) between SNP j and trait t of the given data. With $\gamma_1 = 1$, SPU test weights each SNP equally, and yields the highest power if all the SNPs are associated with the trait t with similar effect sizes and association direction (i.e. all positive or all negative). When the subset of SNPs are associated with the trait t, or their association directions are different, $\mathrm{SPU}(\gamma_1 = 2, \gamma_2)$ is often more powerful. As γ_1 increases, $\mathrm{SPU}(\gamma_1, \gamma_2)$ puts heavier weights on the SNPs which are more strongly associated with the trait t. At the end, as the parameter approaches to ∞ (as an even integer), it only considers the most significant SNP, i.e. $\mathrm{SPU}(\gamma_1 = \infty, \gamma_2) = \sum_{t=1}^{k}\big(\max_{j=1}^{p}|U_{jt}|\big)^{\gamma_2}$.

Similarly, γ_2 controls how much to up-weight the traits that are more likely to be associated with SNPs. $\mathrm{SPU}(\gamma_1, \gamma_2 = 1)$ weights all traits equally and performs best when each trait is equally associated with the SNPs. Similarly, as γ_2 increases, the SPU test over-weights larger trait-based

statistics $S(.;\ t)$; in an extreme case, as $\gamma_2 \to \infty$, we define $\text{SPU}(\gamma_1, \gamma_2 = \infty) = \max_{t=1}^{k} \big|S(\gamma_1;\ t)\big|$. If one is more interested in the most significantly associated single SNP-single trait pair, $\text{SPU}(\gamma_1 = \infty, \gamma_2 = \infty) = \max_{j,t} |U_{jt}|$ can be considered. Using various combinations of γ_1 and γ_2, one can target and fit different association patterns across multiple SNPs and multiple traits, including their varying sparsity levels. As a result, the $\text{SPU}(\gamma_1, \gamma_2)$ tests cover several existing tests as special cases as to be shown.

The aSPUset test chooses (γ_1, γ_2) data-adaptively by taking the minimum p-value of $\text{SPU}(\gamma_1, \gamma_2)$'s as the test statistic for candidates $\gamma_1 \in \Gamma_1$ and $\gamma_2 \in \Gamma_2$,

$$\text{aSPUset} = \min_{\gamma_1, \gamma_2} p_{\gamma_1, \gamma_2}.$$

To assess the significance of all the $\text{SPU}(\gamma_1, \gamma_2)$ and aSPUset test, we use either permutations or MC simulations in a single layer to obtain their p-values. The permutation-based method is useful when the covariance matrix (V) is not easy to estimate (e.g. in a high dimensional setting) or when the usual Normal asymptotics may not hold (e.g. n is not large compared to pk); in contrast, the simulation-based method is more restrictive but computationally more efficient. For the permutation-based method, residual terms $\text{res}_i = Y_i - \widehat{\mu}_i$ in equation (3) are permuted to generate $\text{res}_i^{(b)}$ for $b = 1, ...B$, from which the null score vector $U^{(b)}$ is computed as $U^{(b)} = \sum_{i=1}^{n} X_i' \text{res}_i^{(b)}$. Alternatively, for the simulation method, we simulate the null score vectors independently from the null distribution: $U^{(b)} \sim \mathcal{MN}(0, V)$ for $b = 1, ...B$.

In either case, the null statistics $\text{SPU}(\gamma_1, \gamma_2)^{(b)}$ can be calculated from the null score vectors $U^{(b)}$ for $b = 1, ..., B$. Because all $\text{SPU}(\gamma_1, \gamma_2)$ tests are based on the same null score vectors $U^{(b)}$, we just need to simulate one set of null scores and efficiently compute the null statistics, $\text{SPU}(\gamma_1, \gamma_2)^{(b)}$ tests simultaneously for candidate γ_1, γ_2's. Then the p-value of $\text{SPU}(\gamma_1, \gamma_2)$ is

$$p_{\gamma_1, \gamma_2} = \frac{1 + \sum_{b=1}^{B} (I(|\text{SPU}(\gamma_1, \gamma_2)^{(b)}| \geq |\text{SPU}(\gamma_1, \gamma_2)|)}{B + 1}.$$

We can also simultaneously and efficiently compute the p-value of the aSPUset test based on the same set of the null statistics being used for the SPU tests. Note that for each $\text{SPU}(\gamma_1, \gamma_2)^{(b)}$, we can calculate its p-value as $p_{\gamma_1, \gamma_2}^{(b)} = [\sum_{l \neq b} (I(|\text{SPU}(\gamma_1, \gamma_2)^l| \geq |\text{SPU}(\gamma_1, \gamma_2)^{(b)}|) + 1]/B$. Denote

its minimum as $p^{(b)} = \min_{\gamma_1,\gamma_2} p^{(b)}_{\gamma_1,\gamma_2}$. Then the significance of aSPUset test is obtained as

$$P_{\text{aSPUset}} = \frac{\sum_{b=1}^{B} I(|p^{(b)}| \le |\text{aSPUset}|) + 1}{B+1}.$$

Extensions

As shown by Zhang et al. (2014), in some but not all situations, the GEE-Score test may perform better than the aSPU test for a single SNP and multiple traits; the opposite is true too. Hence, to take advantage of both tests, we combine them by taking their minimum p-value to form a new test statistic,

$$\text{aSPUset-Score} = \min\left(P_{\text{aSPUset}}, P_{\text{Score}}\right). \tag{5}$$

Its p-value can be calculated using simulations or permutations as for aSPUset. The null statistic GEE-Score$^{(b)}$ is obtained from the same score $U^{(b)}$ which is used for SPU$(\gamma_1,\gamma_2)^{(b)}$. Hence the null statistics for SPU$(\gamma_1,\gamma_2)^{(b)}$ and GEE-Score$^{(b)}$ can be computed simultaneously.

We can also consider a variance-weighted version of the SPU and aSPUset tests, called the SPUw and aSPUw-set respectively. Each diagonal element of covariance matrix (V) corresponds to the variance of the individual score element U_{jt}; denote the variance of U_{jt} as V_{jt}. The SPUw test is defined with statistic

$$\text{SPUw}(\gamma_1,\gamma_2) = \sum_{t=1}^{k} \left\{ \left[\sum_{j=1}^{p} (U_{jt}/\sqrt{V_{jt}})^{\gamma_1} \right]^{1/\gamma_1} \right\}^{\gamma_2}.$$

The aSPUw-set test statistic is defined as the one taking the minimum p-value of the multiple SPUw(γ_1,γ_2) tests in the same way as that for aSPUset and SPU(γ_1,γ_2). The SPUw and aSPUw-set tests are invariant to the scale of each trait, and hence may be useful when it is unclear how to standardize multiple traits that are in different scales. However, standardizing the traits (such that their sample variances are all equal to one) may or may not be beneficial; often, the power of the unweighted SPU tests and that of the weighted ones are similar as shown before in other contexts (Pan et al 2014; Zhang et al 2014).

Relationships with other methods

The SPU tests are closely related to some existing tests, covering some as special cases. Guo et al. (2013) proposed a set of nonparametric methods for gene-based multiple trait association analysis, called M-MeanStat, M-MaxStat, and M-TopQ25Stat. Each of the methods of Guo et al. (2013) is built on a generalized Kendall's tau (τ), which quantifies the pairwise association between a single SNP and a single trait. Comparing two sets of statistics: M-MeanStat versus SPUw$(2, 2)$, and M-Max versus SPUw$(\infty, 1)$, we see their equivalence as described in Appendix A.

It is obvious that the SPU(1, 1) test is a burden test, which is optimal if its implicit assumption that each SNP-trait pair is equally associated (with the same association direction) holds. The SPU(2,2) test has connections to several other tests. Zhang et al. (2014) showed that when testing on a single SNP, the SPU(2,2) test under the GEE working independence model is equivalent to MDMR with the Euclidean distance. However, for testing multiple SNPs, the equivalence does not hold (Appendix B). KMR with the linear kernel has the same test statistic as SPU(2,2) if the working correlation matrix R_w of the latter in GEE is correctly specified as the true correlation matrix of Y_i (i.e. $R_w = Corr(Y_i|H_0)$); see Appendix C for derivation. This illustrates the flexibility of our proposed test under GEE, in contrast to the stronger modeling assumption in KMR. Since KMR can be derived based on a random-effects model while the burden test is formulated based on a fixed-effects model, our proposed method can be regarded as combining results from both fixed- and random-effects models.

As to be shown in our numerical studies, the GEE-Score test and MANOVA performed similarly; we establish the equivalence between the GEE-Score test and MANOVA with the Pillai-Bartlett trace (Appendix D). Muller and Peterson (1984) discussed the close relationships among four versions of MANOVA (i.e. with the Pillai-Bartlett trace, Hotelling-Lawley's trace, Wilk's lambda, Roy's largest root), each of which can be written as a function of generalized canonical correlations (CCA). Hence the GEE-Score test is directly related to MANOVA and CCA.

Pathway analysis

We extend the adaptive test for association analysis of a single trait and a pathway (i.e. a set of genes) (Pan et al 2015) to that of multiple traits and a pathway. The main idea is to allow

adaptive weighting at the gene-level, in addition to at the SNP- and trait-levels. Given a pathway S with $|S|$ genes and a single trait t, we partition the score vector according to the genes in S as $U = (U'_{1t}, ..., U'_{|S|,t})'$ with a subvector for gene g (with h_g SNPs) as $U_{gt} = (U_{g,1,t}, ..., U_{g,h_g,t})'$. Denote $\text{SPU}(\gamma_1;\ g,\ t)$ and $\text{SPUpath}(\gamma_1, \gamma_2;\ t)$ as the gene-specific SPU and the pathway-based SPU test statistics for single trait t, respectively. Define a new test statistic $\text{GEE-SPUpath}(\gamma_1, \gamma_2, \gamma_3)$ as the pathway analysis for multiple traits:

$$\text{SPU}(\gamma_1, w_1;\ g,\ t) = \left(\sum_{j=1}^{h_g} (w_{1,g,j} U_{g,j,t})^{\gamma_1} / h_g \right)^{1/\gamma_1},$$

$$\text{SPUpath}(\gamma_1, \gamma_2, w_1, w_2;\ t) = \left(\sum_{g=1}^{|S|} (w_{2,g} \text{SPU}(\gamma_1, w_{1,g};\ g,\ t))^{\gamma_2} \right)^{1/\gamma_2},$$

$$\text{GEE-SPUpath}(\gamma_1, \gamma_2, \gamma_3, w_1, w_2) = \sum_{t=1}^{k} (\text{SPUpath}(\gamma_1, \gamma_2, w_1, w_2;\ t))^{\gamma_3},$$

where the three scalars $\gamma_1, \gamma_2, \gamma_3 > 0$ are specified to control the degrees of weighting the SNPs, genes and traits respectively, $w_1 = (w'_{1,1}, ..., w'_{1,|S|})'$ gives gene-specific weights for the SNPs in gene g as $w_{1,g} = (w_{1,g,1}, ..., w_{1,g,h_g})'$, and $w_2 = (w_{2,1}, ..., w_{2,|S|})'$ gives gene-specific weights for each gene in the pathway S. These weights are specified based on some prior knowledge on the importance of the genes and SNPs; without prior knowledge, we can simply use an equal weight 1 on each gene and each SNP, as used in our later simulations. We employed $\gamma_1 \in \Gamma_1 = \{1, 2, ..., 8\}$ and $\gamma_2, \gamma_3 \in \Gamma_2 = \Gamma_3 = \{1, 2, 4, 8\}$ in later simulations.

Finally, a new adaptive test for pathway analysis, denoted GEE-aSPUpath test, is defined as

$$\text{GEE-aSPUpath} = \min_{\gamma_1 \in \Gamma_1, \gamma_2 \in \Gamma_2, \gamma_3 \in \Gamma_3} p_{\gamma_1, \gamma_2, \gamma_3},$$

where $p_{\gamma_1, \gamma_2, \gamma_3}$ is the p-value of the $\text{GEE-SPUpath}(\gamma_1, \gamma_2, \gamma_3)$ test. The simulation or permutation procedure for generating the null statistics and calculating p-values for all the GEE-SPUpath and GEE-aSPUpath tests are similar to that for the GEE-aSPUset test.

Due to the limited space, we will not discuss the pathway-based tests in the sequel; some simulation results are presented in the Supplementary Materials (File S4).

Results

Real Data Example

ADNI data

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (`adni.loni.usc.edu`). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a 60 million, 5-year public private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials. The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 subjects but ADNI has been followed by ADNI-GO and ADNI-2. To date these three protocols have recruited over 1500 adults, ages 55 to 90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow up duration of each group is specified in the protocols for ADNI-1, ADNI-2 and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see `www.adni-info.org`.

GWAS with ADNI-1 data

One objective of ADNI is to elucidate genetic susceptibility to AD. We conducted a gene-based multi-trait analysis for ADNI-1 data, by using grey matter volumes in the 12 ROIs corresponding to the default mode network (DMN) as intermediate phenotypes. DMN is a network of brain regions

that are active when an individual is at wakeful rest, which includes inferior temporal, medial orbitofrontal, parahippocampal, precuneus and posterior cingulate ROIs (Greicius et al. 2004). Importantly, DMN activity distinguishes cognitively impaired patients such as with Alzheimer's, ADHD, or bipolar disorder from healthy controls (Metin et al. 2015; Meda et al. 2014; Buckner et al. 2008; Greicius et al. 2003, 2004). The grey matter volumetric measures related to the DMN were extracted from the ADNI-1 baseline data.

We included all SNPs with minor allele frequency (MAF) ≥ 0.05, genotyping rate more than 90%, and surviving the Hardy-Weinberg equilibrium test at a significance threshold 0.001. After all rounds of quality control, 519,286 SNPs remained, among which 277,527 SNPs were mapped to 17,557 genes. To consider SNPs in promoter or regulatory regions for each gene, we included SNPs upstream and downstream within 20Kb of each gene. Subjects with more than 10% missing genotypes were excluded, and only non-Hispanic Caucasians whose twelve grey matter volumes in DMN were all measured at baseline were included, resulting in 144 AD patients, 311 MCI subjects, and 180 healthy elderly controls. For covariates, gender, years of education, handedness, age, and intracranial volume (ICV) measured at baseline were included.

To demonstrate the applicability and power of our approach, we applied MANOVA, MDMR (McArdle and Anderson 2001), KMR (Maity et al. 2012), MFLM (Wang et al. 2015) and GEE-based tests, GEE-Score, aSPUset and aSPUset-Score tests. The number of MC simulations or permutations for each method was set $B = 10^3$ at beginning, but was increased up to $B = 10^8$ if an obtained p-value was less than $5/B$, which ensured the identification of the genes at the genome-wide significance level (p-value $< 2.8 \times 10^{-6}$ with a Bonferroni adjustment). When any obtained p-value was less than 1.0e-8, we reported it as 1.0e-8. The p-values of permutation-based aSPUset and of simulation-based aSPUset agreed well (with a Pearson correlation 0.98), thus we reported only permutation-based results. For MFLM, we used beta-smooth basis functions with the Pillai-Bartlett trace as a representative.

The aSPUset and MDMR tests uncovered two loci associated with DMN. Table 1 lists the genes with the highest significance levels. *Genes AMOTL*1 *(on chromosome 11) and APOC1, APOE (on chromosome 19) were identified by both aSPUset and MDMR, but not by other tests, while TOMM40 (on chromosome 19) was only detected by aSPUset. AMOTL1* is known to be involved in cell adhesion and cell signaling (Hamatani et al. 2004). A recent study using a pathway-

enrichment strategy showed that the genes involved in neuronal cell adhesion, and cell signaling are overrepresented in schizophrenia and bipolar disorder (Meda et al. 2014). Anney et al. (2008) identified *AMOTL1* as a gene associated with ADHD. The gene was also highly expressed in thalamus, a brain region implicated in the cognitive impairment of early stage Huntington's disease (Schmouth et al. 2013). Three genes (*APOC1, APOE, TOMM40*) in chromosome 19 could not be readily discerned due to their physical closeness, though their gene sizes (i.e. the numbers of SNPs) varied. The p-values of MDMR became less significant as the gene size increased, while the aSPUset was robust to the number of SNPs. This locus containing *APOE* is well known to be related to Alzheimer's disease and cognitive impairment disorder (Liu et al. 2014; Kamboh et al. 2012; Seshadri et al. 2010).

Table 2 lists the SNPs included in the significant genes. We applied several single SNP-based tests for association with the default mode network. For each method, the permutation or simulation number was increased up to 10^8 to satisfy the genome-wise significance level. As shown in Table 2, none of the SNPs in gene *AMOTL1* was significant, suggesting that a strong association signal was retained only in the gene-level, rather than in the SNP-level. On the other hand, SNP rs429358 contained in three genes (*APOC1, APOE, TOMM40*) was highly significant with p-value of 1.0e-8. *These results lend support for the proposed aSPUset test's potential of being able to recover both multiple weak effects and single strong effects, due to its adaptiveness.*

We explored each identified locus in details in Figures 1 and 2. In Figure 1, a LocusZoom plot (Pruim et al. 2010) illustrates local linkage disequilibrium (LD), recombination patterns and p-values obtained from the single SNP-based aSPU test for DMN. Figure 2 illustrates the association analyses for genes *AMOTL1* and *APOE* respectively. First we obtained p-values from the univariate test between each SNP and each individual trait comprising DMN, then applied SNP-based test (aSPU) between each SNP and DMN (12 traits). Finally, we applied the aSPUset test at the gene level for DMN. The SNPs contained in *AMOTL1* showed strong LD (Figure 1A), and their aggregate effects turned out to be significant at the gene level (Figure 2A). Among the SPU(γ_1, γ_2) tests applied with $\gamma_1, \gamma_2 \in \{1, ..., 8, \infty\}$, SPU(3,2) showed the minimum p-value, implying that weak effects were aggregated for an overall association. In Figure 2B, only one variant (rs429358) in *APOE* was significant, but the significance level of aSPUset did not diminish in the gene level analysis. In testing *APOE*, the p-values of SPU(2,1), SPU(4,1), SPU(6,1), SPU(8,1), and SPU(∞,1)

were tied and the most significant; this suggested that one SNP (rs429358) dominated across in the gene level across all the traits.

Since the proposed test is based on combining all possible single SNP-single trait association pairs, if one would like to identify which pairs contribute most to an overall association, one can simply examine the signifcance levels of the univariate single SNP-single trait association tests. For example, Figure 2 (left panels) illustrates the contribution of each SNP-trait pair for $AMOTL1$ and $APOE$. In the gene $AMOTL1$, the SNP-trait pairs, (rs1367505, R-InferiorTemporal), (rs2033367, R-InferiorTemporal) and (rs333027, L-InferiorParietal), were ranked highest; for $APOE$, the top 3 significant pairs were (rs429358 , R-Precuneus), (rs2075650, L-Precuneus) and (rs429358, L-InferiorParietal).

As shown in Supplementary Materials (File S1), we conducted a single SNP-based GWAS scan for the ADNI-1 data. *Interestingly, no SNP was significant from univariate single SNP–single trait analyses as shown in Figures A and B. Furthermore, only one SNP, rs429358, was significant in single SNP-based multi-trait analyses as shown in Figures C and D. In contrast, two loci (AMOTL1 and APOE) were uncovered by gene-based multi-trait analyses by our proposed new test (Figures E and F).* In all analyses, covariates considered included gender, years of education, handedness, age, and intracranial volume (ICV) measured at baseline. *Taken together, these results clearly demonstrated the advantage and power gain of our proposed gene-based multi-trait analysis.*

Validation with ADNI-GO/2 data

Using the ADNI-1 data as the discovery sample, our GWAS identified two loci associated with DMN. To validate the results, each method was applied to the two genes $AMOTL1$ and $APOE$ using the ADNI-GO/2 data as the validation sample (with $n = 754$). We applied the same SNP-filtering criteria as applied to ADNI-1. Table 3 presents the p-values obtained from each method; no significant association was identified. Due to different genotyping arrays, ADNI-GO/2 data contains different sets of SNPs from those of ADNI-1; we imputed missing SNPs which were originally included in the analysis of ADNI-1, based on the reference samples of HapMap 3 with MaCH (Liu et al. 2013), in order to apply each method to the identical SNP sets of ADNI-1. The aSPUset and aSPUset-Score tests identified gene $APOE$ with p-values 0.019 and 0.024 respectively, which passed the significance threshold 0.05/2 as shown in Table 3, but gene $AMOTL1$ was not significant by

any test. Figure A in Supplementary Materials (File S2) illustrates p-values from single SNP-based testing after adjusting for covariates; SNP rs429358 was associated with DMN (p-value 1.9e-3) by passing the Bonferroni adjusted significance level 0.05/12. Figure B in Supplementary Materials (File S2) presents p-values for the two candidate gene regions based on the ADNI-GO/2 data; the methods include the univariate single SNP–single trait test, the single SNP-based multi-trait aSPU test, and the gene-based multi-trait aSPUset test.

We would mention possible sample differences between ADNI-1 and ADNI-GO/2 cohorts. The ADNI-1 cohort includes three subject groups consisting of 25% AD patients, 50% subjects with MCI (Mild Cognitive Impairment) and 25% CN (Cognitively Normal) subjects; in contrast, the ADN-GO/2 study assigns 754 subjects into five groups: 20% CN , 12% SMC (Significant Memory Concern), 35% EMCI (Early Mild Cognitive Impairment), 17% LMCI (Late Mild Cognitive Impairment), and 16% AD. At least the proportions of the CN subjects and AD patients in the two cohorts are different, which might lead to different association results.

Finally, we combined the two cohorts to form ADNI-1/GO/2 with a larger sample size (about 1400 subjects) and obtained the p-values from the tests for the two candidate gene regions. The two genes were highly significantly associated with the default mode network as shown in Table 3.

Gene-based rare variant analysis of the ADNI sequencing data

The proposed method was applied to analysis of rare variants with the ADNI whole-genome sequencing (WGS) data, consisting of 254 and 500 subjects from ADNI-1 and ADNI-GO/2 respectively. In total, 26,142 genes were included for analyses; all variants inside a gene and those located 25kb of upstream and downstream of the gene were mapped to the gene. Five covariates were adjusted: gender, years of education, handedness, age and ICV. Due to the low frequency of rare variants, the asymptotic assumption for some tests may not hold; we modified each method to avoid using asymptotics. For MANOVA, rather using the usual F-distribution, we permuted residuals (under the null model) to estimate its null distribution; for aSPUset and MFLM, similarly the permutation-based method was applied. We included all rare variants within each gene region; the number of variants within each region ranged from 3 to 750. Sometimes permutation-based MANOVA suffered from rank deficiency when constructing the test statistic and could not be applied to about 600 genes; MFLM also failed for some genes due to rank deficiency.

First we included only rare variants (with MAF < 0.01), then both rare and low-frequency variants (with MAF < 0.05). No gene passed the genome-wide Bonferroni-adjusted significance threshold of 2.8×10^{-6}. The results for each set of rare variants are illustrated in Figures A and B in Supplementary Materials (File S3). MFLM was problematic with an inflation factor around 1.5 in both analyses.

Given that two gene regions were significantly associated with DMN in the previous GWAS analysis, it would be of interest to see whether the rare variants in the two genes were associated. Table 4 reports the p-values for the two candidate genes. No significant associations were detected. Figure C in Supplementary Materials (File S3) depicts the p-values from single trait-based tests, including SKAT, SKAT-O, T1 (a burden test for rare variants with MAF < 0.01), T5 (a burden test for rare and low-frequency variants with MAF < 0.05), minP, and aSPU tests (Wu et al. 2011; Pan et al 2014). T1 and T5 are equivalent to the SPU(1) test with MAF threshold 0.01 and 0.05 respectively. The minP test is similar to the SPU(∞) test.

Simulations

Simulation set-ups

We evaluated the performance of our method along with several existing methods in simulation studies. The simulated data mimicked the association structures for the two genes (*AMOTL1* on chromosome 11 and *TOMM40* on chromosome 19) and default mode network (DMN) in ADNI-1 data. Two factors were considered: association effect size (Set-up 1) and sparsity of association patterns (Set-up 2). For Set-up 1, various effect sizes were created by scaling the regression coefficient estimates obtained from a multivariate linear model (MLM) fitted to the original data. On each gene, an MLM was fitted to the ADNI-1 data, including the covariates (z_i), SNPs (x_i) and DMN (Y_i). For covariates, we included gender, education, handedness, age, and ICV as in the original data analysis. Denote the parameter estimates in an MLM as follows: G_0 is a vector for intercepts; $G = (g_{jt})$ is a $p \times k$ matrix, in which g_{jt} represents the effect size of SNP j on trait t; the element h_{qt} in matrix $H = (h_{qt})$ stands for the qth covariate effect on the tth trait; Σ is the covariance estimate for the multivariate error term. To maintain the true correlation structures among genotype scores $x_i = (x_{i1}, ..., x_{ip})'$ and five covariates $z_i = (z_{i1}, ..., z_{i5})'$, we sampled pairs

(x_i, z_i) from the ADNI-1 data in each simulation. The multiple traits for subject i were generated from a multivariate normal distribution:

$$Y_i \sim \mathcal{MN}(G_0 + \phi \cdot G' x_i + H' z_i, \ \ \Sigma). \tag{6}$$

Here ϕ was a scaling parameter controlling the effect sizes of the SNPs (x_i): with $\phi = 0$, the null hypothesis held and Type I error rates were evaluated; at $\phi = 1$, the effect sizes were set to be equal to the estimated ones from the ADNI-1 data.

For Set-up 2, we varied the sparsity level of the association structure. At a fixed $\phi = 0.5$, we increased the gene size by adding some null SNPs to gene $AMOTL1$. For the null SNPs, the genotype data adjacent to $AMOTL1$ were used. As before, (x_i, z_i) pairs were sampled from the ADNI-1 data. Throughout simulations, 10000 replicates were used for each set-up and the tests were conducted at the significance level $\alpha = 0.05$.

Type I error and power

All the tests showed Type I error rates controlled under the nominal level 0.05 (Table 5). Of note, MDMR resulted in conservative Type I error rates. In Set-up 1 (Table 5), as the association effect size (ϕ) decreased, the aSPUset and aSPUset-Score tests were more powerful than other tests, suggesting the potential usefulness of the proposed tests in identifying causal SNPs with weak effects. Since MFLM was proposed to reduce the dimensionality of the SNP data, it might not be desirable to use MFLM here; it might perform better with larger numbers of SNPs.

In Set-up 2 (Table 6), the aSPUset and aSPUset-Score yielded higher power than other tests as the proportion of the null SNPs in the SNP set increased. Throughout the simulations, the GEE-Score test performed similarly to MANOVA, confirming their equivalence.

Computational time

We reported computational requirement of each method in Table 7 by taking the average computation time for simulation Set-up 2. MANOVA was computationally most efficient, followed by MFLM. As the number of SNPs increased, GEE-Score test and aSPUset-Score test became computationally more demanding, but still feasible.

Conclusions

We have presented a highly adaptive association test for multiple traits and multiple genetic variants. From the GWAS analyses of the ADNI-1 data (File S1 in Supplementary Materials), we observed its potential power gains in identifying cumulative weak effects of multiple associated SNPs in gene $AMOTL1$ with multiple traits, which were undetectable by several other gene-based tests and single SNP-based tests. Given that most common variants have only weak effects for complex diseases and traits, developing testing strategies to improve power in identifying multiple SNPs with weak effects is very important. Our proposed method is developed along this direction. Furthermore, due to its adaptiveness, it also retains power in the presence of only one or few associated SNPs (or traits), as shown for the $APOE$ gene with the ADNI-1 data (while several existing gene-based tests failed to capture). Our proposed adaptive test is in contrast to most of the existing tests, which may be powerful in one or more situations, but not across a wide range of situations. In practice, since the true association pattern for a given gene and traits is unknown, it is unclear which non-adaptive test should be used; it will be convenient and promising to apply an adaptive test such as our proposed one.

We emphasize the potential power gain with the use of multiple traits, especially of intermediate phenotypes for a complex disease such as AD (Chen et al. 2015; Mukherjee et al. 2014). However, since it is unknown how many of, and in what association patterns, the multiple traits are associated with a gene (or a set of SNPs), a straightforward use of any multivariate test may lose, not gain, power. Again, the availability of a powerful and adaptive test such as our proposed one will largely facilitate its easy and effective use in practice.

Finally, we summarize the use of our proposed tests and make some recommendations. To assess an overall association between a set of SNPs and a set of traits, we would recommend the use of the p-value of the aSPUset test. If it is significant, one can check the individual p-values of the SPU(γ_1, γ_2) tests to shed some light on the underlying association pattern. If a larger γ_1 (or γ_2) leads to a more signifcant p-value of the SPU test, it would suggest a more sparse association pattern; that is, perhaps one a fewer number of the SNPs (or traits) are associated. Furthermore, one can examine the p-value from the univariate test for each SNP-trait pair to identify which SNP-trait pairs contribute most to the overall association. For choosing candidate values of γ_1

and γ_2, based on our limited experience, we suggest using $\Gamma_1 = \Gamma_2 = \{1, 2, ..., 8, \infty\}$ by default, though an optimal choice depends on the situation; using a too large or too small set Γ_1 or Γ_2 will lead to loss of power. A general guidance, taking Γ_1 as an example (and similarly for Γ_2), is to use $\Gamma_1 = \{1, 2, ..., C_1, \infty\}$ such that the SPU(C_1, γ_2) test gives a p-value almost equal to that of SPU(∞, γ_2); a larger number of SNPs may require a larger value of C_1. In addition, if some large univariate associations between various SNP-trait pairs are likely to be in opposite directions, only even integers are needed in Γ_1 and Γ_2; if it is known a priori that large univariate associations are mainly in one direction, then using only odd integers may be most powerful; otherwise, both even and odd integers should be used. Given the relationships among the tests, we recommend the use of our proposed aSPUset and aSPUset-Score tests, though MFLM may also perform well for large genes; further evaluations are needed.

Supplementary Materials

The R code for the proposed tests and simulations is available under the Paper Information link at the Genetics website. An R package `GEEaSPU` is to be uploaded to CRAN.

Acknowledgment

The authors are grateful to the reviewers for constructive comments. This research was supported by NIH grants R01GM113250, R01HL105397, and R01HL116720, and by the Minnesota Supercomputing Institute. J.K. was supported by a UMII MnDRIVE fellowship.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimers Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research &

Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Rev December 5, 2013 Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Appendix

Without loss of generality we center both $Y_i = (y_{i1}, y_{i2}, ..., y_{ik})'$ and $x_i = (x_{i1}, x_{i2}, ..., x_{ip})'$ to have their sample means $\sum_{i=1}^{n} Y_i/n = 0$ and $\sum_{i=1}^{n} x_i/n = 0$. We consider the case without covariates, since several methods are only applicable to the case without covariates.

We rewrite data format as a design matrix. Denote Λ as $n \times p$ matrix each row contains subject i's genotype $x_i = (x_{i1}, ..., x_{ip})'$ and Θ as $n \times k$ matrix each row of which consists of multiple traits $Y_i = (y_{i1}, ..., y_{ik})'$. Multivariate analysis can be derived form partitioning of the total sum of squares and cross products (SSCP) matrix, the inner product $\Theta'\Theta$. According to the multivariate linear model, $\Theta = \Lambda \mathrm{B} + \mathrm{E}$, where B is the matrix of model parameters, E is the matrix of errors, the fitted value matrix is defined as $\widehat{\Theta} = \Lambda\widehat{\mathrm{B}} = \Lambda(\Lambda'\Lambda)^{-1}\Lambda'\Theta = \mathrm{H}\Theta$ and the matrix of residuals is $\mathrm{R} = \Theta - \widehat{\Theta} = (\mathrm{I} - \mathrm{H})\Theta$, where H is a hat matrix.

We define each covariance estimate as follows. $S_x = \frac{1}{n}\Lambda'\Lambda$ is a $p \times p$ covariance estimate for genotype scores $x_i = (x_{i1}, ..., x_{ip})'$, and $S_y = \frac{1}{n}\Theta'\Theta$ is a $k \times k$ covariance estimate among k multiple traits $Y_i = (y_{i1}, ..., y_{ik})'$. $S_{yx} = \frac{1}{n}\Theta'\Lambda$ and $S_{xy} = \frac{1}{n}\Lambda'\Theta$ are covariance estimate between two sets of variable x_i and Y_i.

tr(A) stands for sum of diagonal elements of a matrix A. vec(A) represents a linear transformation which converts the matrix (A) into a column vector.

Appendix A SPUw(2, 2) and M-MeanStat; SPUw(∞, 1) and M-Max

For each trait t and SNP j, their pairwise association is quantified by $\tau_{jt} = \sum_{i=1}^{n} x_{ij}(y_{it} - \bar{y}_t) = \sum_{1}^{n} x_{ij}y_{it}$, which follows a normal distribution asymptotically with mean zero and variance $\mathrm{var}(\tau_{jt}|y_t) = \sum_{i=1}^{n} \mathrm{var}(x_{ij})y_{it}^2$ under the null hypothesis. Guo et al. (2015) defined the generalized Kendall's tau statistic, $T_{jt} = \tau_{jt}^2 \mathrm{var}(\tau_{jt}|y_t)^{-1} \sim \chi_1^2$. Based on this, Guo et al. (2013) proposed M-MeanStat and M-MaxStat;

$$\begin{aligned}
\text{M-MeanStat} &= \frac{1}{p}\sum_{t=1}^{k}\sum_{j=1}^{p} T_{jt} \propto \sum_{t=1}^{k}\sum_{j=1}^{p} \frac{(\sum_{i=1}^{n} x_{ij}y_{it})^2}{\sum_{i=1}^{n}\mathrm{var}(x_{ij})y_{it}^2} \approx \sum_{t=1}^{k}\sum_{j=1}^{p}\left(\frac{\sum_{i=1}^{n} x_{ij}y_{it}}{\sqrt{\sum_{i=1}^{n} x_{ij}^2 y_{it}^2}}\right)^2, \\
\text{M-MaxStat} &= \sum_{t=1}^{k}\max_{j=1}^{p} T_{jt} = \sum_{t=1}^{k}\max_{j=1}^{p} \frac{(\sum_{i=1}^{n} x_{ij}y_{it})^2}{\sum_{i=1}^{n}\mathrm{var}(x_{ij})y_{it}^2} \approx \sum_{t=1}^{k}\max_{j=1}^{p}\left(\frac{\sum_{i=1}^{n} x_{ij}y_{it}}{\sqrt{\sum_{i=1}^{n} x_{ij}^2 y_{it}^2}}\right)^2. \qquad (7)
\end{aligned}$$

If a canonical link function and a working independence model are used in GEE, the test statistics of SPUw$(2, 2)$ and SPUw$(\infty, 1)$ are defined by

$$\begin{aligned}
\mathrm{SPUw}(2,2) &\propto \sum_{t=1}^{k}\sum_{j=1}^{p}\left(\frac{\sum_{i=1}^{n} x_{ij}y_{it}}{\sqrt{\sum_{i=1}^{n} x_{ij}^2 \mathrm{var}(y_{it})}}\right)^2 \approx \sum_{t=1}^{k}\sum_{j=1}^{p}\left(\frac{\sum_{i=1}^{n} x_{ij}y_{it}}{\sqrt{\sum_{i=1}^{n} x_{ij}^2 y_{it}^2}}\right)^2, \\
\mathrm{SPUw}(\infty,1) &\propto \sum_{t=1}^{k}\max_{j=1}^{p}\left|\frac{\sum_{i=1}^{n} x_{ij}y_{it}}{\sqrt{\sum_{i=1}^{n} x_{ij}^2 \mathrm{var}(y_{it})}}\right| \approx \sum_{t=1}^{k}\max_{j=1}^{p}\left(\frac{\sum_{i=1}^{n} x_{ij}y_{it}}{\sqrt{\sum_{i=1}^{n} x_{ij}^2 y_{it}^2}}\right)^2. \qquad (8)
\end{aligned}$$

Comparing the two sets of statistics in (7) and (8), we see that M-MeanStat and SPUw$(2, 2)$, and M-Max and SPUw$(\infty, 1)$ are approximately equivalent respectively.

Appendix B SPU(2,2) and MDMR

Under the working independence model, the test statistic of SPU(2,2) is stated as

$$\mathrm{SPU}(2,2) = \sum_{t=1}^{k}\sum_{j=1}^{p}\big(\sum_{i=1}^{n} x_{ij}y_{it}\big)^2 = \mathrm{tr}\big(\Lambda'\Theta\Theta'\Lambda\big) \qquad (9)$$

MDMR (Multivariate Distance Matrix Regression) is a nonparametric modification of traditional Fisher's MANOVA (McArdle and Anderson, 2001). Wessel and Schork (2006) and Zapala and Schork (2012) introduced the method to applications in genetics and genomics. For single trait, it

is closely related to kernel methods (Schaid et al. 2005; Pan 2011).

Suppose d_{ij} represents the distance between subject i and j; let $A = (a_{ij}) = (-1/2\ d_{ij}^2)$ and G its centered version. An F-statistic can be constructed to test the hypothesis that the p regressor variables have no relationship to variation in the distance or dissimilarity of the n subjects reflected in the $n \times n$ distance/dissimilarity matrix. The psuedo F-statistics of MDMR is defined by

$$F = \frac{\mathrm{tr(HGH)}}{\mathrm{tr(I-H)G(I-H)}}$$

If the Euclidean distance (i.e. L_2-norm) is used to construct the distance matrix G= $\Theta\Theta'$, the MDMR test statistic is defined as

$$\mathrm{MDMR} \propto \frac{\mathrm{tr(H\Theta\Theta'H)}}{\mathrm{tr(I-H)\Theta\Theta'(I-H)}} \propto \frac{1}{\mathrm{tr}\big(\mathrm{R}'\mathrm{R}\big)/\mathrm{tr}\big(\widehat{\Theta}'\widehat{\Theta}\big)} \propto \frac{1}{[\mathrm{tr}\big(\widehat{\Theta}'\widehat{\Theta}\big) + \mathrm{tr}\big(\mathrm{R}'\mathrm{R}\big)]/\mathrm{tr}\big(\widehat{\Theta}'\widehat{\Theta}\big)} = \frac{\mathrm{tr}\big(\widehat{\Theta}'\widehat{\Theta}\big)}{\mathrm{tr}\big(\Theta'\Theta\big)}$$

As usual, permutations are used to calculate p-values. Then $\mathrm{tr}\big(\Theta'\Theta\big)$ is invariant across all permutations and can be ignored (Pan, 2011). The test statistic arrives at

$$\mathrm{MDMR} \propto \mathrm{tr}\big(\widehat{\Theta}'\widehat{\Theta}\big) = \mathrm{tr}\big(\Theta'\Lambda(\Lambda'\Lambda)^{-1}\Lambda'\Theta\big) = \mathrm{tr}\big((\Lambda'\Lambda)^{-1}\Lambda'\Theta\Theta'\Lambda\big) \tag{10}$$

If we have a single SNP to be tested, i.e. Λ is an $n \times 1$ matrix; the test statistic (10) reduces to $\mathrm{MDMR} \propto m^{-1}\mathrm{tr}(\Lambda'\Theta\Theta'\Lambda\big) \propto \mathrm{tr}(\Lambda'\Theta\Theta'\Lambda\big)$ with $\Lambda'\Lambda = m$. Hence, SPU(2, 2) and MDMR are equivalent for a single SNP and multiple traits, as established by Zhang et al (2014). However, for multiple SNPs and multiple traits, by comparing (9) and (10), we see that in general they are not equivalent.

Appendix C SPU(2,2) and KMR

With a working correlation matrix R_w in GEE, the SPU(2,2) test can be rewritten as

$$\mathrm{SPU}(2,2) = \mathrm{tr}\big(\Lambda'\Theta R_w^{-1}R_w^{-1}\Theta'\Lambda\big) = \mathrm{tr}(R_w^{-1}\Theta'\Lambda\Lambda'\Theta R_w^{-1}). \tag{11}$$

Maity et al. (2012) introduced multivariate phenotype association analysis by SNP set- or gene-based kernel machine regression (KMR). The authors assumed that the phenotypes are correlated

while the individuals are independent. Suppose $\Psi = (\psi_{pq})$ is the true correlation matrix for k traits with $p = 1, ..., k$, and $q = 1, ..., k$. Define $V_0 = \Psi \otimes I_{n\times n}$, and a kernel matrix $\mathcal{K}_{nk\times nk}$. The score test under the null for KMR (Maity et al. 2012) is defined by

$$\text{KMR} = \text{vec}(\Theta)^{'} V_0^{-1}\, \mathcal{K}\, V_0^{-1}\text{vec}(\Theta) = \text{vec}(\Theta)^{'} V_0^{-1}\, \text{diag}(K_1, ..., K_k)\, V_0^{-1}\text{vec}(\Theta)$$

where each $K_1, ..., K_k$ is an $n \times n$ kernel matrix for each trait. Applying a linear kernel $K_1 =, ..., = K_k = \Lambda\Lambda^{'}$ yields

$$\begin{aligned}\text{KMR} &= \text{vec}(\Theta)^{'} V_0^{-1}\big(I_{k\times k} \otimes \Lambda\Lambda^{'}\big)V_0^{-1}\text{vec}(\Theta) = \text{vec}(\Theta\Psi^{-1})^{'}\big(I \otimes \Lambda\Lambda^{'}\big)\text{vec}(\Theta\Psi^{-1}) \\ &= \text{vec}(\Theta\Psi^{-1})^{'}\text{vec}(\Lambda\Lambda^{'}\Theta\Psi^{-1}) = \text{tr}(\Psi^{-1}\Theta^{'}\Lambda\Lambda^{'}\Theta\Psi^{-1}). \qquad (12)\end{aligned}$$

KMR (12) has the same test statistic as the GEE-SPU(2) test (11) if the working correlation R_w is the true correlation structure of Y_i (i.e. $\Psi = R_w = Corr(Y_i|H_0)$).

Appendix D GEE-Score test and MANOVA

The GEE-Score test statistic with a working independence model in GEE is

$$\begin{aligned}\text{GEE-Score} &= \text{vec}(\Lambda^{'}\Theta)^{'}\big(S_y \otimes nS_x\big)^{-1}\text{vec}(\Lambda^{'}\Theta) = n\ \text{vec}(S_{xy})^{'}\big(S_y^{-1} \otimes S_x^{-1}\big)\text{vec}(S_{xy}) \\ &= n\ \text{tr}\big(S_y^{-1}S_{yx}S_x^{-1}S_{xy}\big).\end{aligned}$$

In MANOVA, a measure of the strength of association between Θ (multiple traits) and Λ (genotype scores) for the multivariate model $\Theta = \Lambda\text{B} + \text{E}$ depends on a partition of matrix of total SSCP i.e. $\Theta^{'}\Theta = \widehat{\Theta}^{'}\widehat{\Theta} + \text{R}^{'}\text{R}$ (Haase, 2011). Considering the Pillai-Bartlett (PB) trace, the MANOVA test statistic is stated as $\text{tr}\big(\widehat{\Theta}^{'}\widehat{\Theta}(\Theta^{'}\Theta)^{-1}\big) = \text{tr}\big(\Theta^{'}\Lambda(\Lambda^{'}\Lambda)^{-1}\Lambda^{'}\Theta(\Theta^{'}\Theta)^{-1}\big)$, which can be written in an alternate form $\text{tr}\big(S_{yx}S_x^{-1}S_{xy}S_y^{-1}\big) = \text{tr}\big(S_y^{-1}S_{yx}S_x^{-1}S_{xy}\big)$. Hence, the GEE-Score test and MANOVA using the PB trace are equivalent.

References

Alzheimer's Association. (2015a). Alzheimer's disease facts and figures. *Alzheimer's & Dementia* **11**, 332-384.

Alzheimer's Association. (2015b). Changing the trajectory of Alzheimer's disease: How a treatment by 2025 saves lives and dollars. Available at `http://www.alz.org/documents_custom/trajectory.pdf`.

Anney, RJ., Lasky-Su J., O'Dúshláine, C., Kenny, E., Neale, BM. Mulligan, A., Franke, B., Zhou, K., Chen, W., Christiansen, H., et al. (2008) Conduct disorder and ADHD: evaluation of conduct problems as a categorical and quantitative trait in the international multicentre ADHD genetics study. *Am J Med Genet B Neuropsychiatr Genet* **147B** (8), 1369–1378.

Aschard, H., Vilhjalmsson, B., Wu, C., Greliche, N., Morange, PE., Wolpin, B., Tregouet, DA., Kraft, P. (2014) Maximizing the power in principal components analysis of correlated phenotypes. *Am J Hum Genet* **94** (5), 662–676.

Balthazar, M., Weiler, M., Campos, B., Rezende, T., Damasceno, B., Cendes, F. (2014) Alzheimer as a Default Mode Network Disease: A grey matter, functional and structural connectivity study. *Neurology* **83** (10) P6.324.

Buckner, RL., Andrews-Hanna, JR., Schacter, DL. (2008) The brain's default network: anatomy, function, and relevance to disease. *Ann N Y Acad Sci* **1124**, 1–38.

Chen, CH., Peng, Q., Schork, AJ., Lo, MT., Fan, CC., Wang, Y, Desikan, RS. et al. (2015) Large-scale genomics unveil polygenic architecture of human cortical surface area. *Nat Commun* **6**, 7549. doi: 10.1038/ncomms8549.

Damoiseaux, JS., Seeley, WW., Zhou, J., Shirer, WR., Coppola, G., Karydas, A., Rosen, HJ., Miller, BL., Kramer, JH., Greicius, MD.; Alzheimer?s Disease Neuroimaging Initiative (2012). Gender modulates the APOE $\varepsilon 4$ effect in healthy older adults: convergent evidence from functional brain connectivity and spinal fluid tau levels. *J Neurosci* **32**, 8254–8262.

Ferreira, MA., Purcell, SM. (2009). A multivariate test of association. *Bioinformatics* **25**, 132–133.

Glahn, DC., Winkler, AM., Kochunov, P., Almasy, L., Duggirala, R., Carless, MA., Curran, JC., Olvera, RL, .Laird, AR. Smith, SM., et al. (2010) Genetic control over the resting brain *Proc Natl Acad Sci* **107** (3), 1223–1228.

Greicius, MD., Srivastava, G., Reiss, AL., Menon, V. (2004) Default mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proc Natl Acad Sci* **101**, 4637–4642.

Guo, X., Liu, Z., Wang, X., Zhang, H. (2013) Genetic association test for multiple traits at gene level. *Genet Epidemiol* **37** (1), 122–129.

Jones, L., Holmans, PA., Hamshere, ML., Harold, D., Moskvina, V., Ivanov, D., Pocklington, A., Abraham, R., Hollingworth, P., Sims, R., et al. (2010) Genetic evidence implicates the immune system and cholesterol metabolism in the aetiology of Alzheimer's disease. *PLoS ONE* **5**, e13950.

Jones, DT., Machulda, MM., Vemuri, P., McDade, EM., Zeng, G., Senjem, ML., Gunter, JL., Przybelski, SA., Avula, RT., Knopman, DS., Boeve, BF., Petersen, RC., Jack, CR. Jr. (2011) Age-related changes in the default mode network are more advanced in Alzheimer disease. *Neurology* **77** (16), 1524-1531.

Haase, RF. Multivariate General Linear Models. (pp. 59-103). Thousand Oaks, CA: SAGE Publications; (2011).

He, Y., Chen, Z., Gong, GL., Evans, A. (2009) Neuronal networks in Alzheimer's disease. *Neuroscientist* **15**, 333–350.

Hong, MG., Reynolds, CA., Feldman, AL., Kallin, M., Lambert, JC., Amouyel, P., Ingelsson, E., Pedersen, NL., Prince, JA. (2012) Genome-wide and gene-based association implicates FRMD6 in Alzheimer disease. *Hum Mutat* **33**, 521–529.

Kamboh, MI., Demirci, FY., Wang, X., Minster, RL., Carrasquillo, MM., Pankratz, VS., Younkin, SG., Saykin, AJ., Alzheimer's Disease Neuroimaging Initiative, Jun, G., Baldwin, C., Logue, MW., Buros, J., Farrer, L., Pericak-Vance, MA,, Haines, JL., Sweet, RA., Ganguli, M.,

727 Feingold, E., Dekosky, ST., Lopez, OL., Barmada, MM. (2012) Genome-wide association
728 study of Alzheimer's disease. *Transl Psychiatry* **15**;2:e117. doi: 10.1038/tp.2012.45.

729 Karch, CM., Cruchaga, C., Goate, AM. (2014) Alzheimer's disease genetics: from the bench to
730 the clinic. *Neuron* **83** (1), 11–26.

731 Klei, L., Luca, D., Devlin, B., Roeder, K. (2008) Pleiotropy and principal components of heri-
732 tability combine to increase power for association analysis. *Genet Epidemiol* **32**, 9-19.

733 Liang, K., Zeger, S. (1986) Longitudinal data analysis using generalized linear models. *Biometrika*
734 **73**, 13–22.

735 Lin, J., Zhu, H.T., Knickmeyer, R., Styner, M., Gilmore, J. H., Ibrahim, J.G. (2012) Projection
736 regression models for multivariate imaging phenotype. *Genetic Epidemiology* **36**, 631-641.

737 Liu, D., Lin, X., Ghosh, D. (2007) Semiparametric regression of multidimensional genetic pathway
738 data: least-squares kernel machines and linear mixed models. *Biometrics* **63**, 1079–1088.

739 Liu, EY., Li, M., Wang, W., Li, Y. (2013). MaCH-Admix: Genotype Imputation for Admixed
740 Populations. *Genet Epidemiol* **37** (1):25–37.

741 Liu, G., Yaoc, L., Liu, J., Jiang, Y., Ma, G., the Genetic and Environmental Risk for Alzheimer's
742 disease (GERAD1) Consortium, Chen, Z., Zhao, B., Li, K. (2014) Cardiovascular disease
743 contributes to Alzheimer's disease: evidence from large-scale genome-wide association studies.
744 *Neurobiol of Aging* **35** (4), 786–792.

745 Maity, A., Sullivan, PF., Tzeng, JY. (2012) Multivariate phenotype association analysis by marker-
746 set kernel machine regression. *Genet Epidemiol* **36**, 686–695.

747 Manolio, TA., Collins, FS., Cox, NJ., Goldstein, DB., Hindorff, L.A., Hunter, DJ., McCarthy,
748 MI., Ramos, EM., Cardon, LR., Chakravarti, A., et al. (2009) Finding the missing herita-
749 bility of complex diseases. *Nature* **461**, 747–753.

750 Marei, H., Althani, A., El Zowalaty, M.,, Albanna, MA.,, Cenciarelli, C., Wang, T., Caceci, T.
751 (2015) Common and Rare Variants Associated with Alzheimer's Disease. *J Cell Physiol* doi:
752 10.1002/jcp.25225.

McArdle, BH., Anderson, MJ. (2001) Fitting multivariate models to community data: A comment on distance-based redundancy analysis. *Ecology* **82**, 290–297.

Metin, B., Krebs, RM., Wiersema, JR., Verguts, T., Gasthuys, R., van der Meere, JJ., Achten, E., Roeyers, H., Sonuga-Barke, E. (2015) Dysfunctional modulation of default mode network activity in attention-deficit/hyperactivity disorder. *J Abnorm Psychol* **124** (1), 208–214.

Meda, SA., Ruao, G., Windemuth, A., O'Neil, K., Berwise, C, Dunn, SM., Boccaccio LE., Narayanan B., Kocherla, M., Sprooten, E, Keshavan, MS., Tamminga CA., Sweeney JA., Clementz, BA., Calhoun, VD, Pearlson, GD. (2014) Multivariate analysis reveals genetic associations of the resting default mode network in psychotic bipolar disorder and schizophrenia. *Proc Natl Acad Sci* **111** (19), E2066–2075.

Mukherjee, S., Kim, S., Ramanan, VK., Gibbons, LE., Nho, K., Glymour, MM., Ertekin-Taner, N., Montine, TJ., Saykin, AJ., Crane, PK., the Alzheimer's Disease Neuroimaging Initiative. (2014) Gene-based GWAS and biological pathway analysis of the resilience of executive functioning. *Brain Imaging Behav* **8**, 110–118.

Muller, KE., Peterson, BL. (1984) Practical methods for computing power in testing the multivariate general linear hypothesis. *Comput Stat Data An* **2**, 143–158.

Pan, W. (2011) Relationship between genomic distance-based regression and kernel machine regression for multi-marker association testing. *Genet Epidemiol* **35** (4), 211–216.

Pan, W., Kim, J., Zhang, Y., Shen, X. and Wei, P. (2014) A powerful and adaptive association test for rare variants. *Genetics* **197**, 1081–1095.

Pan, W., Kwak, I., Wei, P. (2015) A powerful pathway-based adaptive test for genetic association with common or rare variants. *Am. J. Hum. Genet.* **97**, 86-98.

Pruim, RJ., Welch, RP., Sanna, S., Teslovich, TM., Chines, PS., Gliedt, TP., Boehnke, M., Abecasis, GR., Willer, CJ. (2010) LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337.

Ridge, PG., Mukherjee, S., Crane, PK., Kauwe, JS., Alzheimer's Disease Genetics Consortium (2013). Alzheimer's disease: analyzing the missing heritability. *PLoS ONE* **8**, e79771.

780 Saykin, A.J., Shen, L., Yao, X., Kim, S., Nho, K., Risacher, S. L., ..., Alzheimer's Disease Neu-
781 roimaging Initiative. (2015). Genetic studies of quantitative MCI and AD phenotypes in
782 ADNI: Progress, opportunities, and plans. *Alzheimer's & Dementia* **11**, 792-814.

783 Schifano, ED., Li, L., Christiani, DC., Lin, X. (2013) Genome-wide association analysis for mul-
784 tiple continuous secondary phenotypes. *Am J Hum Genet* **92**, 744-759.

785 Schmouth, JF., Castellarin, M., Laprise, S., Banks, KG., Bonaguro, RJ., McInerny, SC., and
786 others. (2013) Non-coding-regulatory regions of human brain genes delineated by bacterial
787 artificial chromosome knock-in mice. *BMC biol* **11**, 106.

788 Shen, L., Thompson, PM., Potkin, SG., Bertram, L., Farrer, LA., Foroud, TM., for the Alzheimer's
789 Disease Neuroimaging Initiative. (2014) Genetic analysis of quantitative phenotypes in AD
790 and MCI: imaging, cognition and biomarkers. *Brain Imaging Behav* **8** (2), 183–207.

791 Sherva, R., Tripodis, Y., Bennett, DA., Chibnik, LB., Crane, PK., de Jager, PL., Farrer, LA.,
792 Saykin, AJ., Shulman, JM., Naj, A., et al.; GENAROAD Consortium; Alzheimer's Disease
793 Neuroimaging Initiative; Alzheimer's Disease Genetics Consortium (2014). Genome-wide
794 association study of the rate of cognitive decline in Alzheimer's disease. *Alzheimer's Dement*
795 **10**, 45–52.

796 Schaid, DJ., McDonnell, SK., Hebbring, SJ., Cunningham, JM., Thibodeau, SN. (2005) Non-
797 parametric tests of association of multiple genes with human disease. *Am J Hum Genet* **76**,
798 780-793.

799 Shen, L., Kim, S., Risachera, SL., Nho, K., Swaminathan, S., Westa, JD., Foroudd, T., et al.
800 (2010) Whole genome association study of brain-wide imaging phenotypes for identifying
801 quantitative trait loci in MCI and AD: A study of the ADNI cohort. *NeuroImage* **53**, 1051–
802 1063.

803 Tang, CS., Ferreira, MAR. (2012) A gene-based test of association using canonical correlation
804 analysis. *Bioinformatics* **28** (6), 845–850. doi:10.1093/bioinformatics/bts051.

805 Tzeng, JY., Zhang, D., Pongpanich, M., Smith, C., McCarthy, MI., Sale, MM., Worrall, BB.,
806 Hsu, FC., Thomas, DC., Sullivan, PF. (2011) Studying gene and gene-environment effects of

807 uncommon and common variants on continuous traits: a marker-set approach using gene-trait
808 similarity regression. *Am J Hum Genet* **89**, 277–288.

809 Van der Sluis, S., Dolan, CV., Li, J., Song, Y., Sham, P., Posthuma, D., Li, M. (2015) MGAS: a
810 powerful tool for multivariate gene-based genome-wide association analysis. *Bioinformatics*
811 **31** (7), 1007–1015.

812 Wang, K., Abbott, D. (2007) A principal components regression approach to multilocus genetic
813 association studies. *Genetic Epidemiology* **32**, 108-118.

814 Wang, X., Lee, S., Zhu, X., Redline, S. Lin, X. (2013) GEE-based SNP set association test for
815 continuous and discrete traits in family-based association studies. *Genet Epidemiol* **37**, 778–
816 786.

817 Wang, Y., Liu, A., Mills, JL., Boehnke, M., Wilson, AF., Bailey-Wilson, JE., Xiong, M., Wu,
818 CO., Fan, R. (2015) Pleiotropy analysis of quantitative traits at gene level by multivariate
819 functional linear models. *Genet Epidemiol* **39** (4), 259–75.

820 Wessel, J., Schork, NJ. (2006) Generalized genomic distance-based regression methodology for
821 multilocus association analysis. *Am J Hum Genet* **79**, 792–806.

822 Wu, MC., Lee, S., Cai, T., Li, Y., Boehnke, M., and Lin, X. (2011) Rare Variant Associa- tion
823 Testing for Sequencing Data Using the Sequence Kernel Association Test (SKAT). *Am J Hum*
824 *Genet* **89**, 82–93.

825 Zapala, MA., Schork, NJ. (2012) Statistical properties of multivariate distance matrix regression
826 for high-dimensional data analysis. *Front Genet* **3**, 190.

827 Zhang, Y., Xu, X., Shen, X., Pan, W., for the Alzheimer's Disease Neuroimaging Initiative.
828 (2014) Testing for association with multiple traits in generalized estimation equations, with
829 application to neuroimaging data. *NeuroImage* **96**, 309–325.

Table 1: P-values of the gene-based association tests for DMN with the ADNI-1 data.

					GEE						
Gene-region	#SNPs	Chr	Position		Score	aSPUset	aSPUset-Score	MANOVA	MDMR	KMR	MFLM
AMOTL1	6	11	94121155	94269566	1.18e0-4	**1.0e-08**	**1.0e-08**	7.73e-05	**3.48e-07**	0.451	7.73e-05
APOC1	4	19	50089760	50134446	6.14e-04	**1.0e-08**	**1.0e-08**	3.45e-04	**4.42e-08**	0.342	2.30e-04
APOE	6	19	50080878	50124490	1.27e-03	**1.0e-08**	**1.0e-08**	7.93e-04	**2.21e-07**	0.268	5.97e-04
TOMM40	10	19	50066316	50118786	0.023	**1.0e-08**	**1.0e-08**	1.86e-02	6.99e-06	0.569	1.04e-03

Table 2: P-values of the single SNP-based association tests for DMN for the significant gene-regions (±20kb) with the ADNI-1 data.

					GEE					
Gene	Chr	aSPUset	SNP	Position	Score	SPU(2)	SPU(∞)	aSPU	MANOVA	MDMR
AMOTL1	11	1.0e-08	rs1367505	94186285	8.0e-05	2.4e-07	2.8e-05	5.1e-07	5.1e-05	2.1e-07
			rs10501816	94187396	0.417	0.151	0.237	0.158	0.432	0.186
			rs2033367	94195356	1.2e-04	8.0e-07	6.5e-05	1.6e-06	9.1e-05	3.01e-07
			rs2241667	94203379	8.0e-04	1.6e-06	1.3e-04	3.9e-06	1.8e-04	8.0e-06
			rs333027	94225561	5.0e-04	1.6e-05	9.5e-05	3.1e-05	4.6e-04	6.9e-05
			rs333025	94227040	0.02	0.025	0.030	0.045	0.015	0.022
APOC1	19	1.0e-08	rs8106922	50093506	0.236	0.116	0.212	0.183	0.244	0.128
			rs405509	50100676	0.420	0.156	0.207	0.186	0.422	0.184
			rs439401	50106291	7.0e-04	2.3e-06	1.2e-05	3.1e-06	4.1e-04	2.2e-05
			rs429358	50103781	1.0e-05	**4e-0**8	8.3e-06	**1.0e-08**	2.1e-06	**1.25e-08**
APOE	19	1.0e-08	rs157580	50087106	3.1e-03	1.4e-04	8.8e-04	9.0e-05	3.1e-03	3.9e-4
			rs2075650	50087459	9.0e-04	3.8e-06	2.2e-03	1.2e-06	2.9e-04	1.5e-05
			rs8106922	50093506	0.236	0.116	0.212	0.183	0.244	0.128
			rs405509	50100676	0.420	0.156	0.207	0.186	0.422	0.184
			rs439401	50106291	7.0e-04	2.3e-06	1.2e-05	3.1e-06	4.1e-04	2.2e-05
			rs429358	50103781	1.0e-05	**4e-0**8	8.3e-06	**1.0e-08**	2.1e-06	**1.25e-08**
TOMM40	19	1.0e-08	rs2075642	50069307	0.842	0.711	0.471	0.629	0.840	0.662
			rs387976	50070900	0.073	0.031	0.036	0.040	0.068	0.067
			rs11667640	50071631	0.262	0.034	0.012	0.021	0.265	0.035
			rs6859	50073874	0.728	0.076	0.299	0.057	0.729	0.072
			rs157580	50087106	3.1e-03	1.4e-04	8.8e-04	9.0e-05	3.1e-03	3.9e-4
			rs2075650	50087459	9.0e-04	3.8e-06	2.2e-03	1.2e-06	2.9e-04	1.5e-05
			rs8106922	50093506	0.236	0.116	0.212	0.183	0.244	0.128
			rs405509	50100676	0.420	0.156	0.207	0.186	0.422	0.184
			rs439401	50106291	7.0e-04	2.3e-06	1.2e-05	3.1e-06	4.1e-04	2.2e-05
			rs429358	50103781	1.0e-05	**4e-0**8	8.3e-06	**1.0e-08**	2.1e-06	**1.25e-08**

Table 3: P-values of the gene-based association tests with the ADNI-GO/2 and ADNI-1/GO/2 data.

						GEE					
Data	Gene-region	#SNPs	Chr	Position		Score	aSPUset	aSPUset-Score	MANOVA	MDMR	MFLM
ADNI-GO/2	AMOTL1	13	11	94481507	94629918	0.723	0.896	0.940	0.698	0.716	0.638
	APOE	13	19	45389277	45432652	0.083	0.042	0.056	0.097	0.366	0.974
ADNI-GO/2 with	AMOTL1	6	11	-		0.639	0.552	0.576	0.638	0.918	0.638
identical SNP sets of ADNI-1	APOE	6	19	-		0.308	**0.019**	**0.024**	0.292	0.065	0.292
ADNI-1/GO/2 with	AMOTL1	6	11	-		**1.0e-08**	**1.0e-08**	**1.0e-08**	**1.0e-08**	**1.0e-08**	**1.0e-08**
identical SNP sets of ADNI-1	APOE	6	19	-		**1.0e-08**	**1.0e-08**	**4.45e-06**	**1.0e-08**	**1.0e-08**	**4.45e-06**

Table 4: P-values of the gene-based tests for rare variant–DMN association with the ADNI sequencing data.

Filtering criteria	Gene-region	# SNPs	Chr	Position		aSPUset	MANOVA	MFLM
MAF$<$ 0.05	AMOTL1	536	11	94481507	94629918	0.298	0.176	0.148
	APOE	153	19	45389277	45432652	0.104	0.837	0.476
MAF$<$ 0.01	AMOTL1	265	11	94481507	94629918	0.835	0.193	0.151
	APOE	84	19	45389277	45432652	0.874	0.833	0.189

Table 5: Simulation setup 1: Type I errors ($\phi = 0$) and power ($\phi \neq 0$) under varying genetic effect sizes.

AMOTL1 (6 SNPs)

	GEE						
ϕ	Score	SPU(2,2)	aSPUset	aSPUset-Score	MANOVA	MDMR	MFLM
0	0.0479	0.0528	0.0530	0.0522	0.0490	0.0353	0.0490
0.2	0.1078	0.1837	0.1659	0.1654	0.1128	0.0964	0.1128
0.3	0.2325	0.3494	0.3159	0.3328	0.2394	0.2135	0.2394
0.4	0.4657	0.5571	0.5079	0.5559	0.4764	0.4130	0.4764
0.5	0.7436	0.7614	0.7156	0.7967	0.7528	0.6607	0.7528
0.6	0.9288	0.9008	0.8722	0.9452	0.9341	0.8608	0.9341
0.7	0.9913	0.9677	0.9550	0.9926	0.9921	0.9611	0.9921

TOMM40 (10 SNPs)

	GEE						
ϕ	Score	SPU(2,2)	aSPUset	aSPUset-Score	MANOVA	MDMR	MFLM
0	0.0488	0.0483	0.0482	0.0495	0.0505	0.0323	0.0532
0.2	0.1051	0.1719	0.1347	0.1369	0.1110	0.0903	0.1116
0.3	0.2177	0.3643	0.2763	0.2889	0.2262	0.2053	0.2169
0.4	0.4429	0.6121	0.5018	0.5330	0.4605	0.4246	0.4256
0.5	0.5800	0.7304	0.6231	0.6673	0.5958	0.5593	0.5664
0.6	0.7196	0.8271	0.7369	0.7904	0.7346	0.6885	0.7036
0.7	0.8405	0.8983	0.8293	0.8856	0.8489	0.8015	0.8231

Table 6: Simulation setup 2: power under varying sparsity levels of association pattern.

AMOTL1+ Null SNPs

# total	# causal	# null	GEE					
SNPs	SNPs	SNPs	Score	aSPUset	aSPUset-Score	MANOVA	MDMR	MFLM
6	6	0	0.7436	0.7156	0.7967	0.7528	0.6607	0.7528
12	6	6	0.5332	0.6495	0.6923	0.5427	0.4904	0.5228
18	6	12	0.4160	0.6149	0.6336	0.4291	0.3884	0.3882
30	6	24	0.2950	0.4495	0.4617	0.3055	0.2819	0.2872
60	6	54	0.1813	0.3120	0.3150	0.1981	0.1756	0.2124
80	6	74	0.1442	0.2912	0.2912	0.1661	0.1434	0.1697

Table 7: Mean computing times (in seconds) for simulation setup 2.

# total	GEE					
SNPs	Score	aSPUset	aSPUset-Score	MANOVA	MDMR	MFLM
12	1.1597	1.2472	1.6261	0.0149	24.2924	0.0354
18	1.3398	1.5062	2.2552	0.0156	22.2903	0.0385
30	2.2541	1.8766	3.7482	0.0172	21.5940	0.0449
60	6.5183	2.8785	11.1315	0.0211	19.3995	0.0612
80	11.8868	3.5546	20.4237	0.0243	18.4600	0.0722

Figure 1: LocusZoom for two loci identified by aSPUset and MDMR: LD structure in each locus and p-values obtained from the single SNP-based aSPU test are presented.

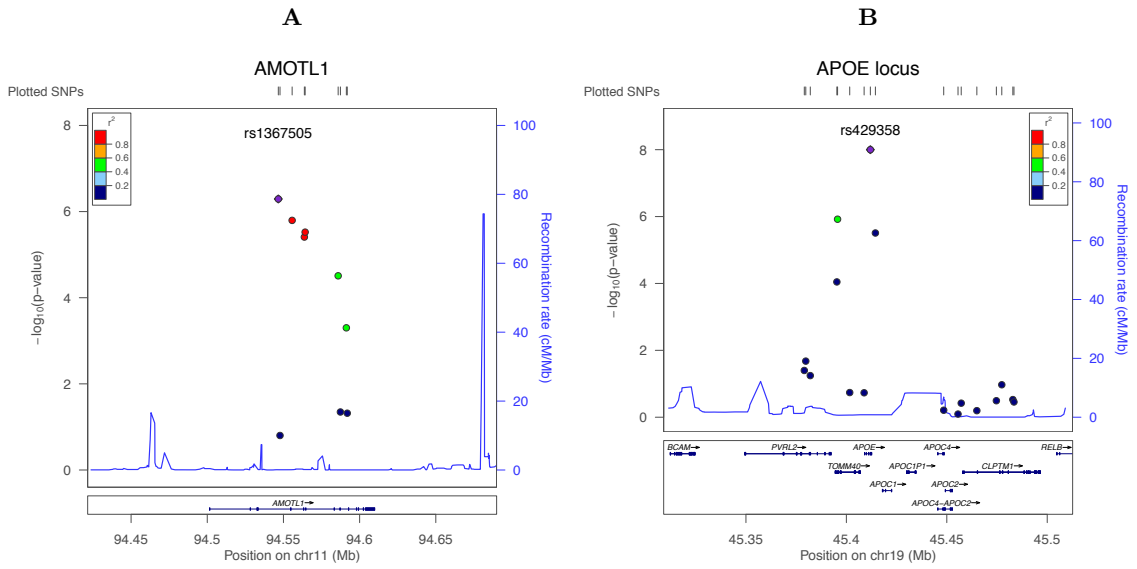
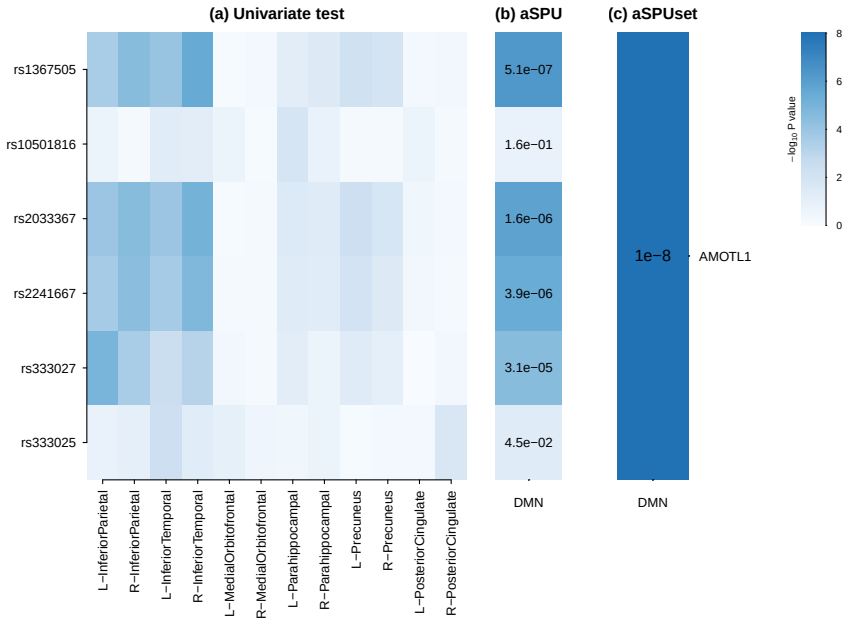


Figure 2: P-values of the association tests for DMN and SNPs for genes *AMOTL1* and *APOE*: (a) univariate test for single SNP–single trait association; (b) aSPU test for single SNP–multitrait association; (c) aSPUset for gene–multitrait association.

A AMOTL1



B APOE