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

JUNGHI KIM¹, YIWEI ZHANG¹, WEI PAN¹,

FOR THE ALZHEIMER’S DISEASE NEUROIMAGING INITIATIVE²

¹Division 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
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
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 $i = 1, \ldots, n$, we observe $k$ traits $Y_i = (y_{i1}, \ldots, y_{ik})'$, $q$ covariates $z_i = (z_{i1}, \ldots, z_{iq})'$ and a set of single nucleotide polymorphisms (SNPs) $x_i = (x_{i1}, \ldots, x_{ip})'$, 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

$$g(\mu_i) = Z_i \varphi + X_i \beta = H_i \theta$$

with $H_i = (Z_i X_i)$, parameters $\theta = (\varphi', \beta')'$, and a link function $g(.)$. The regression coefficients $\beta = (\beta_{11}, \ldots, \beta_{p1}, \ldots, \beta_{1k}, \ldots, \beta_{pk})'$ is a $pk \times 1$ vector, in which $\beta_{jt}$ represents the effect of the $j$th SNP on the $t$th trait, while the element $\varphi_{st}$ of $\varphi = (\varphi_{11}, \ldots, \varphi_{(q+1)1}, \ldots, \varphi_{1k}, \ldots, \varphi_{(q+1)k})'$ is the effect size of the $s$th covariate on the $t$th trait. Liang and Zeger (1986) proposed estimating $\varphi$ and $\beta$ by solving the generalized estimating equations (GEE):

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

with $D_i = \partial \mu_i / \partial \theta'$ and $V_i = \phi A_i^{1/2} R_w(\alpha) A_i^{1/2}$, where $\phi$ is a dispersion parameter, $A_i = \text{diag}\{v(\mu_{i1}), \ldots, v(\mu_{ik})\}$ models the variances with a variance function $v(\mu_i)$, and $R_w(\alpha)$ is a working correlation matrix with possibly some unknown parameters $\alpha$. 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 $U_\theta$ and its variance-covariance matrix $\text{Cov}(U_\theta)$ are

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

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

The covariance matrix can be partitioned according to the score components for $\varphi$ and $\beta$: $\text{Cov}(U_\theta) = (V_{11} V_{12}) (V_{21} V_{22})$. 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 $\varphi$ 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 $\varphi$, denoted as $\widehat{\varphi}$, is the solution to the generalized score equation $U_{\varphi, \beta=0} = \sum_{i=1}^{n} Z_i'(Y_i - \hat{\mu}_i) = 0$. The marginal mean is estimated by $\hat{\mu}_i = g(Z_i \widehat{\varphi})^{-1}$.

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

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

$U_\beta$ asymptotically follows a multivariate normal distribution $\mathcal{MN}(0, \hat{\Sigma}_\beta)$ under $H_0$, where $\hat{\Sigma}_\beta = V_{22} - V_{21} V_{11}^{-1} V_{12}$. $U_\beta$ 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 $\beta_{jt}$ 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

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

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 $U^{(b)}_\beta \sim \mathcal{MN}(0, \hat{\Sigma}_\beta)$ for $b = 1, ..., B$ and
obtain the null statistics $\text{GEE-Score}^{(b)} = U^{(b)}_{\beta} \tilde{\Sigma}_{\beta}^{-1} U^{(b)}_{\beta}$. 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 $\beta$ 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 $\Gamma$; e.g. $\Gamma = \{1, 2, ..., 8, \infty\}$. The statistical power of an SPU(\gamma) test depends on the choice of $\gamma \in \Gamma$. When $\gamma$ is an odd integer, the SPU(\gamma) 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 $\gamma$, the SPU(\gamma) 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 $\gamma$ 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 $U^{(b)} = (U^{(b)}_{j1}, ..., U^{(b)}_{jk})'$ is generated from $\mathcal{MN}(0, V)$, from which the null statistics $\text{SPU}(\gamma)^{(b)} = \sum_{t=1}^{k} (U^{(b)}_{jt})^\gamma$ can be obtained for each $\gamma$. Then the p-value can be
calculated as $p_\gamma = \left[ \sum_{b=1}^{B} I(SPU(\gamma) \leq SPU(\gamma)^{(b)}) + 1 \right] / (B + 1)$.

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

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

where $p_\gamma$ is p-value of SPU(\gamma). By MC simulations (or permutations), the p-value of aSPU, along with those of all SPU(\gamma) 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 random-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](http://www4.stat.ncsu.edu/~maity/software.html) and [https://www.nih.gov/about/org/diphr/bbb/software/fan/Pages/default.aspx](https://www.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 $aSPU_{set}$, 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 $U = (U_{11}, \ldots, U_{p1}, \ldots, U_{1k}, \ldots, U_{pk})'$ 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, $\gamma_1$ and $\gamma_2$, to control the degrees of weighting over the SNPs and over the traits respectively, and the two parameters are chosen data-adaptively. A larger $\gamma_1$ puts more weights on the SNPs more likely to be associated with a given trait, while a larger $\gamma_2$ 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 $\text{SPU}(\gamma_1, \gamma_2)$ combines the single trait-based statistics:

$$S(\gamma_1; t) = \left( \sum_{j=1}^{p} (U_{jt})^{\gamma_1} \right)^{1/\gamma_1}, \quad \text{SPU}(\gamma_1, \gamma_2) = \sum_{t=1}^{k} (S(\gamma_1; t))^{\gamma_2}. \quad (4)$$

Here candidate integers $\gamma_1 \geq 1$ and $\gamma_2 \geq 1$ are to be chosen from two pre-selected parameter sets $\Gamma_1$ and $\Gamma_2$. 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, $\text{SPU}(\gamma_1 = 2, \gamma_2)$ is often more powerful. As $\gamma_1$ increases, $\text{SPU}(\gamma_1, \gamma_2)$ puts heavier weights on the SNPs which are more strongly associated with the trait $t$. At the end, as the parameter approaches to $\infty$ (as an even integer), it only considers the most significant SNP, i.e. $\text{SPU}(\gamma_1 = \infty, \gamma_2) = \sum_{t=1}^{k} \left( \max_{j=1}^{p} |U_{jt}| \right)^{\gamma_2}$.

Similarly, $\gamma_2$ controls how much to up-weight the traits that are more likely to be associated with SNPs. $\text{SPU}(\gamma_1, \gamma_2 = 1)$ weights all traits equally and performs best when each trait is equally associated with the SNPs. Similarly, as $\gamma_2$ increases, the SPU test over-weights larger trait-based
statistics \( S(; t) \); in an extreme case, as \( \gamma_2 \to \infty \), we define \( \text{SPU}(\gamma_1, \gamma_2 = \infty) = \max_{t=1}^k |S(\gamma_1; t)| \). 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 \( \gamma_1 \) and \( \gamma_2 \), 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 \((\gamma_1, \gamma_2)\) 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 \( \text{res}_i = Y_i - \hat{\mu}_i \) in equation (3) are permuted to generate \( \text{res}_i^{(b)} \) for \( b = 1, \ldots B \), 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 \( b = 1, \ldots B \).

In either case, the null statistics \( \text{SPU}(\gamma_1, \gamma_2)^{(b)} \) can be calculated from the null score vectors \( U^{(b)} \) for \( b = 1, \ldots, B \). 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 \( \gamma_1, \gamma_2 \)'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 \( p_{\gamma_1, \gamma_2}^{(b)} = \frac{\sum_{t \neq b} (I(|\text{SPU}(\gamma_1, \gamma_2)^{(t)}| \geq |\text{SPU}(\gamma_1, \gamma_2)^{(b)}|) + 1)}{B} \). 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

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

**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,

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

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 $\text{SPU(}\gamma_1, \gamma_2)^{(b)}$. Hence the null statistics for $\text{SPU(}\gamma_1, \gamma_2)^{(b)}$ 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$^{(\gamma_1, \gamma_2)}$ tests in the same way as that for aSPUset and SPU$^{(\gamma_1, \gamma_2)}$. 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 ($\tau$), 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 
$U = (U_{1,t}',\ldots,U_{|S|,t}')'$ with a subvector for gene $g$ (with $h_g$ SNPs) as $U_{g,t} = (U_{g,1,t},\ldots,U_{g,h_g,t})'$. Denote 
$SPU(\gamma_1; g, t)$ and $SPUpath(\gamma_1, \gamma_2; t)$ as the gene-specific SPU and the pathway-based SPU test 
statistics for single trait $t$, respectively. Define a new test statistic $GEE-SPUpath(\gamma_1, \gamma_2, \gamma_3)$ as the 
pathway analysis for multiple traits:

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

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

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

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, $w_1 = (w_{1,1}',\ldots,w_{1,|S|}')'$ gives gene-specific weights for the SNPs in gene 
g as $w_{1,g} = (w_{1,g,1},\ldots,w_{1,g,h_g})'$, and $w_2 = (w_{2,1},\ldots,w_{2,|S|})'$ 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 $\gamma_1 \in \Gamma_1 = \{1,2,\ldots,8\}$ 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

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

where $p_{\gamma_1, \gamma_2, \gamma_3}$ is the p-value of the $GEE-SPUpath(\gamma_1, \gamma_2, \gamma_3)$ 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 AMOTL1 (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($\gamma_1, \gamma_2$) tests applied with $\gamma_1, \gamma_2 \in \{1, \ldots, 8, \infty\}$, 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(\infty,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 significance 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
ADNI-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 Im-
pairment), 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 sequenc-
ing (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 asymp-
totics. 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 $2.8 \times 10^{-6}$. 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 $q$th covariate effect on the $t$th trait; $\Sigma$ 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:

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

(6)

Here \(\phi\) 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 \((\phi)\) decreased, the aSPUs and aSPUs-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 aSPUs and aSPUs-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 aSPUs-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(γ1, γ2) tests to shed some light on the underlying association pattern. If a larger γ1 (or γ2) leads to a more significant 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 γ1
and $\gamma_2$, 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 $\Gamma_1$ or $\Gamma_2$ will lead to loss of power. A general guidance, taking $\Gamma_1$ as an example (and similarly for $\Gamma_2$), is to use $\Gamma_1 = \{1, 2, ..., C_1, \infty\}$ such that the SPU($C_1, \gamma_2$) test gives a p-value almost equal to that of SPU($\infty, \gamma_2$); 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 $\Gamma_1$ and $\Gamma_2$; 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 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 $\Lambda$ as $n \times p$ matrix each row contains subject $i$’s genotype $x_i = (x_{i1}, ..., x_{ip})'$ and $\Theta$ 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, $\Theta = \Lambda B + E$, where $B$ is the matrix of model parameters, $E$ is the matrix of errors, the fitted value matrix is defined as $\hat{\Theta} = \Lambda \hat{B} = \Lambda (\Lambda' \Lambda)^{-1} \Lambda' \Theta = H \Theta$ and the matrix of residuals is $R = \Theta - \hat{\Theta} = (I - H) \Theta$, 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 $	au_{jt} = \sum_{i=1}^{n} x_{ij}(y_{it} - \bar{y}_t) = \sum_{i=1}^{n} x_{ij}y_{it}$, which follows a normal distribution asymptotically with mean zero and variance $\text{var}(\tau_{jt}|y_t) = \sum_{i=1}^{n} \text{var}(x_{ij})y_{it}^2$ under the null hypothesis. Guo et al. (2015) defined the generalized Kendall’s tau statistic, $T_{jt} = \tau_{jt}^2 \text{var}(\tau_{jt}|y_t)^{-1} \sim \chi^2_1$. Based on this, Guo et al. (2013) proposed M-MeanStat and M-MaxStat;

$$
\text{M-MeanStat} = k \sum_{t=1}^{p} \sum_{j=1}^{p} T_{jt} \propto k \sum_{t=1}^{p} \sum_{j=1}^{p} \left( \sum_{i=1}^{n} x_{ij}y_{it} \right)^2 \approx k \sum_{t=1}^{p} \sum_{j=1}^{p} \left( \sum_{i=1}^{n} x_{ij}y_{it} \right)^2 \text{var}(x_{ij})y_{it},
$$

$$
\text{M-MaxStat} = k \sum_{t=1}^{p} \max_{j=1}^{p} T_{jt} = k \sum_{t=1}^{p} \max_{j=1}^{p} \left( \sum_{i=1}^{n} x_{ij}y_{it} \right)^2 \approx k \sum_{t=1}^{p} \max_{j=1}^{p} \left( \sum_{i=1}^{n} x_{ij}y_{it} \right)^2 \text{var}(x_{ij})y_{it}.
$$

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

$$
\text{SPUw}(2, 2) \propto k \sum_{t=1}^{p} \sum_{j=1}^{p} \left( \sum_{i=1}^{n} \frac{x_{ij}y_{it}}{\sqrt{\sum_{i=1}^{n} x_{ij}^2 \text{var}(y_{it})}} \right)^2 \approx k \sum_{t=1}^{p} \sum_{j=1}^{p} \left( \sum_{i=1}^{n} \frac{x_{ij}y_{it}}{\sqrt{\sum_{i=1}^{n} x_{ij}^2 y_{it}^2}} \right)^2,
$$

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

Comparing the two sets of statistics in (7) and (8), we see that M-MeanStat and SPUw(2, 2), and M-Max and SPUw(∞, 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

$$
\text{SPU}(2, 2) = \sum_{t=1}^{k} \sum_{j=1}^{p} \left( \sum_{i=1}^{n} x_{ij}y_{it} \right)^2 = \text{tr}(\Lambda^T \Theta \Theta^T \Lambda) \quad (9)
$$

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 (Schaed 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 pseudo F-statistics of MDMR is defined by

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

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

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

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

$$\text{MDMR} \propto \text{tr}(\hat{\Theta}' \hat{\Theta}) = \text{tr}(\Theta' \Lambda (\Lambda' \Lambda)^{-1} \Lambda' \Theta) = \text{tr}((\Lambda' \Lambda)^{-1} \Lambda' \Theta \Theta' \Lambda)$$ (10)

If we have a single SNP to be tested, i.e. $\Lambda$ is an $n \times 1$ matrix; the test statistic (10) reduces to

$$\text{MDMR} \propto m^{-1} \text{tr}(\Lambda' \Theta \Theta' \Lambda) \propto \text{tr}(\Lambda' \Theta \Theta' \Lambda)$$

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

$$\text{SPU}(2,2) = \text{tr}(\Lambda' \Theta R_w^{-1} R_w^{-1} \Theta' \Lambda') = \text{tr}(R_w^{-1} \Theta' \Lambda \Lambda' \Theta R_w^{-1})$$. (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 $p = 1, \ldots, k,$ and $q = 1, \ldots, k$. Define $V_0 = \Psi \otimes I_{n \times n}$, and a kernel matrix $K_{nk \times nk}$. The score test under the null for KMR (Maity et al. 2012) is defined by

$$K_{nk \times nk} = \psi^{-1} V_0^{-1} K V_0^{-1} \psi^{-1} \text{vec}(\Theta) = \psi^{-1} V_0^{-1} \text{vec}(\Theta)$$

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

$$K_{nk \times nk} = \psi^{-1} V_0^{-1} (I_{k \times k} \otimes \Lambda \Lambda') V_0^{-1} \psi^{-1} \text{vec}(\Theta) = \psi^{-1} V_0^{-1} \text{vec}(\Theta \Psi^{-1}) (I \otimes \Lambda \Lambda') \text{vec}(\Theta \Psi^{-1})$$

$$= \psi^{-1} V_0^{-1} \text{vec}(\Lambda \Lambda' \Theta \Psi^{-1}) = tr(\psi^{-1} \Theta' \Lambda \Lambda' \Theta \Psi^{-1}).$$  \hspace{1cm} (12)

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. $\Psi = R_w = \text{Corr}(Y_i | H_0)$).

Appendix D  GEE-Score test and MANOVA

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

$$\text{GEE-Score} = \text{vec}(\Lambda' \Theta)' (S_y \otimes nS_x)^{-1} \text{vec}(\Lambda' \Theta) = n \text{vec}(S_{xy})' (S_y^{-1} \otimes S_x^{-1}) \text{vec}(S_{xy})$$

$$= n \text{tr}(S_y^{-1} S_{yx} S_x^{-1} S_{xy}).$$

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


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

tinguishes Alzheimer’s disease from healthy aging: evidence from functional MRI. *Proc Natl


Jones, L., Holmans, PA., Hamshere, ML., Harold, D., Moskvina, V., Ivanov, D., Pocklington,
ONE* **5**, e13950.

Jones, DT., Machulda, MM., Vemuri, P., McDade, EM., Zeng, G., Senjem, ML., Gunter, JL.,
Age-related changes in the default mode network are more advanced in Alzheimer disease.
*Neurology* **77** (16), 1524-1531.

Publications; (2011).


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

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.,


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

<table>
<thead>
<tr>
<th>Gene-region</th>
<th>#SNPs</th>
<th>Chr</th>
<th>Position</th>
<th>Score</th>
<th>aSPUset</th>
<th>aSPUset-Score</th>
<th>MANOVA</th>
<th>MDMR</th>
<th>KMR</th>
<th>MFLM</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMOTL1</td>
<td>6</td>
<td>11</td>
<td>94121155</td>
<td>94289566</td>
<td>1.18e-04</td>
<td>1.0e-08</td>
<td>1.0e-08</td>
<td>7.73e-05</td>
<td>3.48e-07</td>
<td>0.451</td>
</tr>
<tr>
<td>APOC1</td>
<td>4</td>
<td>19</td>
<td>500995760</td>
<td>50134446</td>
<td>6.14e-04</td>
<td>1.0e-08</td>
<td>1.0e-08</td>
<td>3.45e-04</td>
<td>4.42e-08</td>
<td>0.342</td>
</tr>
<tr>
<td>APOE</td>
<td>6</td>
<td>19</td>
<td>50089878</td>
<td>50124490</td>
<td>1.27e-03</td>
<td>1.0e-08</td>
<td>1.0e-08</td>
<td>7.93e-04</td>
<td>2.21e-07</td>
<td>0.268</td>
</tr>
<tr>
<td>TOMM40</td>
<td>10</td>
<td>19</td>
<td>50066316</td>
<td>50118786</td>
<td>0.023</td>
<td>1.0e-08</td>
<td>1.0e-08</td>
<td>1.86e-02</td>
<td>6.99e-06</td>
<td>0.569</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr</th>
<th>aSPUset</th>
<th>SNP</th>
<th>Position</th>
<th>Score</th>
<th>SPU(2)</th>
<th>SPU(∞)</th>
<th>aSPU</th>
<th>MANOVA</th>
<th>MDMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMOTL1</td>
<td>11</td>
<td>1.0e-08</td>
<td>rs1367505</td>
<td>94186285</td>
<td>8.0e-05</td>
<td>2.4e-07</td>
<td>2.8e-05</td>
<td>5.1e-07</td>
<td>5.1e-05</td>
<td>2.1e-07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs10501816</td>
<td>94187396</td>
<td>0.417</td>
<td>0.151</td>
<td>0.237</td>
<td>0.158</td>
<td>0.432</td>
<td>0.186</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs2033367</td>
<td>94195356</td>
<td>1.2e-04</td>
<td>8.0e-07</td>
<td>6.5e-05</td>
<td>1.6e-06</td>
<td>9.1e-05</td>
<td>3.01e-07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs2241667</td>
<td>94203379</td>
<td>8.0e-04</td>
<td>1.6e-06</td>
<td>1.3e-04</td>
<td>3.9e-06</td>
<td>1.8e-04</td>
<td>8.0e-06</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs333027</td>
<td>94225561</td>
<td>5.0e-04</td>
<td>8.0e-07</td>
<td>6.5e-05</td>
<td>1.6e-06</td>
<td>9.1e-05</td>
<td>3.01e-07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs333025</td>
<td>94227040</td>
<td>0.02</td>
<td>0.025</td>
<td>0.030</td>
<td>0.045</td>
<td>0.015</td>
<td>0.022</td>
</tr>
<tr>
<td>APOC1</td>
<td>19</td>
<td>1.0e-08</td>
<td>rs8106922</td>
<td>50093506</td>
<td>0.236</td>
<td>0.116</td>
<td>0.212</td>
<td>0.183</td>
<td>0.244</td>
<td>0.128</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs405509</td>
<td>50100676</td>
<td>0.420</td>
<td>0.156</td>
<td>0.207</td>
<td>0.186</td>
<td>0.422</td>
<td>0.184</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs439401</td>
<td>50106291</td>
<td>7.0e-04</td>
<td>2.3e-06</td>
<td>1.2e-05</td>
<td>3.1e-06</td>
<td>4.1e-04</td>
<td>2.2e-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs429358</td>
<td>50103781</td>
<td>1.0e-05</td>
<td>4e-08</td>
<td>8.3e-06</td>
<td>1.0e-08</td>
<td>2.1e-06</td>
<td>1.25e-08</td>
</tr>
<tr>
<td>APOE</td>
<td>19</td>
<td>1.0e-08</td>
<td>rs157580</td>
<td>50087106</td>
<td>3.1e-03</td>
<td>1.4e-04</td>
<td>8.8e-04</td>
<td>9.0e-05</td>
<td>3.1e-03</td>
<td>3.9e-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs2075650</td>
<td>50087459</td>
<td>9.0e-04</td>
<td>3.8e-06</td>
<td>2.2e-03</td>
<td>1.2e-06</td>
<td>2.9e-04</td>
<td>1.5e-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs8106922</td>
<td>50093506</td>
<td>0.236</td>
<td>0.116</td>
<td>0.212</td>
<td>0.183</td>
<td>0.244</td>
<td>0.128</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs405509</td>
<td>50100676</td>
<td>0.420</td>
<td>0.156</td>
<td>0.207</td>
<td>0.186</td>
<td>0.422</td>
<td>0.184</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs439401</td>
<td>50106291</td>
<td>7.0e-04</td>
<td>2.3e-06</td>
<td>1.2e-05</td>
<td>3.1e-06</td>
<td>4.1e-04</td>
<td>2.2e-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs429358</td>
<td>50103781</td>
<td>1.0e-05</td>
<td>4e-08</td>
<td>8.3e-06</td>
<td>1.0e-08</td>
<td>2.1e-06</td>
<td>1.25e-08</td>
</tr>
<tr>
<td>TOMM40</td>
<td>19</td>
<td>1.0e-08</td>
<td>rs2075642</td>
<td>50069307</td>
<td>0.842</td>
<td>0.711</td>
<td>0.471</td>
<td>0.629</td>
<td>0.840</td>
<td>0.662</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs387976</td>
<td>50070900</td>
<td>0.073</td>
<td>0.031</td>
<td>0.036</td>
<td>0.040</td>
<td>0.068</td>
<td>0.067</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs11667640</td>
<td>50071631</td>
<td>0.262</td>
<td>0.034</td>
<td>0.012</td>
<td>0.021</td>
<td>0.265</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs6859</td>
<td>50073874</td>
<td>0.728</td>
<td>0.076</td>
<td>0.099</td>
<td>0.057</td>
<td>0.729</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs157580</td>
<td>50087106</td>
<td>3.1e-03</td>
<td>1.4e-04</td>
<td>8.8e-04</td>
<td>9.0e-05</td>
<td>3.1e-03</td>
<td>3.9e-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs2075650</td>
<td>50087459</td>
<td>9.0e-04</td>
<td>3.8e-06</td>
<td>2.2e-03</td>
<td>1.2e-06</td>
<td>2.9e-04</td>
<td>1.5e-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs8106922</td>
<td>50093506</td>
<td>0.236</td>
<td>0.116</td>
<td>0.212</td>
<td>0.183</td>
<td>0.244</td>
<td>0.128</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs405509</td>
<td>50100676</td>
<td>0.420</td>
<td>0.156</td>
<td>0.207</td>
<td>0.186</td>
<td>0.422</td>
<td>0.184</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs439401</td>
<td>50106291</td>
<td>7.0e-04</td>
<td>2.3e-06</td>
<td>1.2e-05</td>
<td>3.1e-06</td>
<td>4.1e-04</td>
<td>2.2e-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs429358</td>
<td>50103781</td>
<td>1.0e-05</td>
<td>4e-08</td>
<td>8.3e-06</td>
<td>1.0e-08</td>
<td>2.1e-06</td>
<td>1.25e-08</td>
</tr>
</tbody>
</table>
Table 3: P-values of the gene-based association tests with the ADNI-GO/2 and ADNI-1/GO/2 data.

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

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

<table>
<thead>
<tr>
<th>Filtering criteria</th>
<th>Gene-region</th>
<th># SNPs</th>
<th>Chr</th>
<th>Position</th>
<th>aSPUset</th>
<th>MANOVA</th>
<th>MFLM</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAF &lt; 0.05</td>
<td>AMOTL1</td>
<td>536</td>
<td>11</td>
<td>94481507</td>
<td>0.298</td>
<td>0.176</td>
<td>0.148</td>
</tr>
<tr>
<td></td>
<td>APOE</td>
<td>153</td>
<td>19</td>
<td>45389277</td>
<td>0.104</td>
<td>0.837</td>
<td>0.476</td>
</tr>
<tr>
<td>MAF &lt; 0.01</td>
<td>AMOTL1</td>
<td>265</td>
<td>11</td>
<td>94481507</td>
<td>0.835</td>
<td>0.193</td>
<td>0.151</td>
</tr>
<tr>
<td></td>
<td>APOE</td>
<td>84</td>
<td>19</td>
<td>45389277</td>
<td>0.874</td>
<td>0.833</td>
<td>0.189</td>
</tr>
</tbody>
</table>
Table 5: Simulation setup 1: Type I errors ($\phi = 0$) and power ($\phi \neq 0$) under varying genetic effect sizes.

<table>
<thead>
<tr>
<th></th>
<th>AMOTL1 (6 SNPs)</th>
<th></th>
<th>TOMM40 (10 SNPs)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GEE</td>
<td></td>
<td>GEE</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Score</td>
<td>SPU(2,2)</td>
<td>Score</td>
</tr>
<tr>
<td>0</td>
<td>0.0479</td>
<td>0.0528</td>
<td>0.0488</td>
</tr>
<tr>
<td>0.2</td>
<td>0.1078</td>
<td>0.1837</td>
<td>0.0483</td>
</tr>
<tr>
<td>0.3</td>
<td>0.2325</td>
<td>0.3494</td>
<td>0.1719</td>
</tr>
<tr>
<td>0.4</td>
<td>0.4657</td>
<td>0.5571</td>
<td>0.1347</td>
</tr>
<tr>
<td>0.5</td>
<td>0.7436</td>
<td>0.7614</td>
<td>0.2177</td>
</tr>
<tr>
<td>0.6</td>
<td>0.9288</td>
<td>0.9008</td>
<td>0.4429</td>
</tr>
<tr>
<td>0.7</td>
<td>0.9913</td>
<td>0.9677</td>
<td>0.7196</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>AMOTL1+ Null SNPs</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td># total SNPs</td>
<td># causal SNPs</td>
<td># null SNPs</td>
<td>GEE</td>
</tr>
<tr>
<td># causal SNPs</td>
<td>Score</td>
<td>aSPUset</td>
<td>Score</td>
</tr>
<tr>
<td></td>
<td>aSPUset-Score</td>
<td>MANOVA</td>
<td>MFLM</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.7436</td>
<td>0.7528</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>0.1078</td>
<td>0.5332</td>
</tr>
<tr>
<td>18</td>
<td>6</td>
<td>0.2325</td>
<td>0.4160</td>
</tr>
<tr>
<td>30</td>
<td>6</td>
<td>0.4657</td>
<td>0.4950</td>
</tr>
<tr>
<td>60</td>
<td>6</td>
<td>0.7436</td>
<td>0.2950</td>
</tr>
<tr>
<td>80</td>
<td>6</td>
<td>0.9288</td>
<td>0.1813</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th># total SNPs</th>
<th>GEE</th>
</tr>
</thead>
<tbody>
<tr>
<td># causal SNPs</td>
<td>Score</td>
<td>aSPUset</td>
</tr>
<tr>
<td></td>
<td>aSPUset-Score</td>
<td>MANOVA</td>
</tr>
<tr>
<td></td>
<td>Score</td>
<td>MFLM</td>
</tr>
<tr>
<td>12</td>
<td>1.1597</td>
<td>1.2472</td>
</tr>
<tr>
<td>18</td>
<td>1.3398</td>
<td>1.5062</td>
</tr>
<tr>
<td>30</td>
<td>2.2541</td>
<td>1.8766</td>
</tr>
<tr>
<td>60</td>
<td>5.1834</td>
<td>2.8785</td>
</tr>
<tr>
<td>80</td>
<td>11.8868</td>
<td>3.5546</td>
</tr>
</tbody>
</table>

35
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.

A

AMOTL1

Plotted SNPs

rs1367505

Position on chr11 (Mb)

B

APOE locus

Plotted SNPs

rs429358

Position on chr19 (Mb)
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