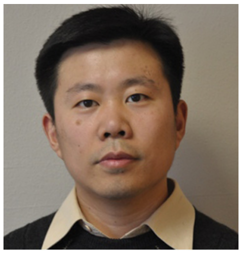
Send Orders for Reprints to reprints@benthamscience.ae

Current Genomics, 2016, 17, 403-415

Detecting Gene-Gene Interactions Associated with Multiple Complex Traits with U-Statistics

Ming Li[1], Changshuai Wei[2], Yalu Wen[3], Tong Wang[4] and Qing Lu[4,5,*]

[1]Department of Epidemiology and Biostatistics, Indiana University at Bloomington, Bloomington, IN 47405, U.S.A; [2]Department of Epidemiology and Biostatistics, University of North Texas Health Science Center, Fort Worth, TX 76107, U.S.A; [3]Department of Statistics, University of Auckland, Auckland 1010, New Zealand; [4]Department of Health Statistics, School of Public Health, Shanxi Medical University, Taiyuan, Shanxi 030001, P.R. China; [5]Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, MI 48824, U.S.A



Q. Lu

Abstract: Many complex diseases, such as psychiatric and behavioral disorders, are commonly characterized through various measurements that reflect physical, behavioral and psychological aspects of diseases. While it remains a great challenge to find a unified measurement to characterize a disease, the available multiple phenotypes can be analyzed jointly in the genetic association study. Simultaneously testing these phenotypes has many advantages, including considering different aspects of the disease in the analysis, and utilizing correlated phenotypes to improve the power of detecting disease-associated variants. Furthermore, complex diseases are likely caused by the interplay of multiple genetic variants through complicated mechanisms. Considering gene-gene interactions in the joint association analysis of complex diseases could further increase our ability to discover genetic variants involving complex disease pathways. In this article, we propose a stepwise U-test for joint association analysis of multiple loci and multiple phenotypes. Through simulations, we demonstrated that testing multiple phenotypes simultaneously could attain higher power than testing one single phenotype at a time, especially when there are shared genes contributing to multiple phenotypes. We also illustrated the proposed method with an application to Nicotine Dependence (ND), using datasets from the Study of Addition, Genetics and Environment (SAGE). The joint analysis of three ND phenotypes identified two SNPs, rs10508649 and rs2491397, and reached a nominal *P*-value of 3.79e-13. The association was further replicated in two independent datasets with *P*-values of 2.37e-05 and 7.46e-05.

Keywords: Pleiotropy, Nicotine dependence, Population-based association studies.

Received: February 11, 2015 **Revised: May 26, 2015** **Accepted: June 06, 2015**

1. INTRODUCTION

Genome-wide association studies (GWASs) have been commonly adopted for investigating the genetic basis of complex human diseases, successfully identifying thousands of single nucleotide polymorphisms (SNPs) associated with complex diseases [1, 2]. However, for many complex diseases, the findings to date only explain a small percentage of heritability [3-5]. The genetic etiology of complex human diseases has remained largely unknown, and detecting genetic variants that account for the "missing heritability" has continued to be a major goal and challenge for the coming decade. The GWASs have commonly used a single-locus approach to test the association between a single SNP and a disease outcome of interest. Such a single-locus and single-phenotype strategy could have limitations on fully utilizing information from the genotype level and the phenotype level. First, complex diseases are usually caused by multiple genetic variants, each conferring a small to moderate effect. The single-locus tests could be under-powered due to the low effect sizes of causal variants and the burden of multiple testing. In addition, genetic variants may interact with one another through complicated mechanisms, and thus, may be overlooked if they are tested separately without considering possible interaction effects. Second, a complex disease may manifest with a wide variety of features, such as multiple measurements of a disease, intermediate phenotypes, subphenotypes, and endophenotypes. These phenotypes may better characterize the underlying disease etiology, and hence, provide more information than a single disease outcome [6]. In genetics, it is also a common phenomenon that shared genetic variants may simultaneously influence multiple phenotypes (i.e., pleiotropy) [7]. The successful identification of shared genetic variants contributing to seemingly distinct phenotypes will help elucidate the common genetic cause of these phenotypes, and will promote the development of a more efficient strategy to treat or prevent these diseases.

*Address correspondence to this author at the Department of Health Statistics, School of Public Health, Shanxi Medical University, Taiyuan, Shanxi 030001, P.R. China; Tel: 517.353.8623 x137; Fax: 517.432.1130; E-mail: qlu@epi.msu.edu

1875-5488/16 $58.00+.00

There is also a growing interest in analyzing multiple related phenotypes in GWAS [6, 8, 9]. For instance, Phenome-Wide Association Studies (PheWAS) are interested in analyzing multiple phenotypes instead of single phenotype. These studies commonly adopt a conventional single-SNP/single-phenotype approach in the analysis. Although convenient, the single-SNP/single-phenotype approach may significantly increase the number of statistical tests, leading to reduced power. To account for the issue of multiple testing due to the increased number of phenotypes, Lange *et al.* proposed to use principal components of phenotypes (PCP) for dimensionality reduction [10]. However, such a strategy is less straight forward for interpretation, because the outcome becomes a linear combination of phenotypes. More importantly, a PCP captures key phenotype information but not necessarily phenotype information related to genetic information. To address this limitation, Klei *et al.* extended the PCP method with a principle component of heritability (PCH) method [11]. However, this PCH method required estimating a PCH for each single SNP, which is computationally intensive for high-dimensional data. A number of other methods were also proposed by using the generalized estimation equations (GEE) and the generalized Kendall's Tau test [12, 13]. It has been shown that these multi-phenotype tests have improved performance over a single-phenotype test. However, these available methods are mainly developed to test each single SNP at a time. Statistical methods that consider the joint effect of multiple variants with multiple phenotypes are still under-developed.

During the past decade, multi-locus tests considering gene-gene interactions have been increasingly used in genetic association analyses [14-20]. Non-parametric methods, such as U-statistic-based methods, have shown great promise for high-dimensional data analysis, especially when the underlying phenotype distributions and modes of inheritance are unknown. Various formations of U-Statistics have been adopted for multi-locus association tests [21-24]. For example, Schaid *et al.* proposed a U-statistic-based score test that summarized a set of SNPs, and then examined their joint association with a phenotype [21]. Wei *et al.* extended this method by using data-adaptive weights for different genetic variants [22]. We and others have further considered possible interactions among genetic variants, and proposed a forward-U test and a likelihood ratio Mann-Whitey test for quantitative phenotypes and binary phenotypes, respectively [23, 24]. These multi-locus methods have emerged as promising tools in the joint association analysis of a single disease phenotype. It is also of great interest to extend those methods for the analyses of multiple phenotypes.

In this article, we propose a U-Statistic-based method, a stepwise U-test, for testing the joint association between multiple genetic variants and multiple phenotypes. It can be viewed as an extension of a recently developed forward U-test for single-phenotype analyses [23]. The proposed method has the following properties: 1) it searches forwardly for SNPs that are associated with one or more phenotypes; 2) it filters backwardly to remove phenotypes that are not relevant to genetic variants; and 3) it tests the joint effect among SNPs while allowing for possible interactions. Through simulations, we have shown the proposed method had improved performance over a single-phenotype test. We also illustrated the proposed method with an application to Nicotine Dependence (ND).

2. METHODS

Suppose we have a study population of *N* subjects. Each subject has *T* measured phenotypes, and is genotyped with *K* SNPs. Let $Y_i = (y_{i,1} ...... y_{i,T})$ and $X_i = (x_{i,1} ...... x_{i,K})$ be the phenotypes and SNP genotypes for subject *i*. Here, we assume that all phenotypes are quantitative and may have unknown distributions. We further assume 1) a subset of phenotypes is associated with part of *K* SNPs; 2) a subset of SNPs influences part of *T* measured phenotypes with possible interactions.

2.1. U-Statistic

We have recently proposed a U-Statistic-based method, referred to as forward U-test, to test the joint association analysis between multiple loci and a single phenotype [23]. In this article, we extend forward U-test for testing the joint association between multiple loci and multiple phenotypes. Following the similar notations, we assume *k* disease-associated SNPs comprising *L* multi-locus genotypes, denoted by G_1, G_2, . . . , G_L. The selection process of *k* disease-associated SNPs is detailed below (Section 2.2). Here, a multi-locus genotype, G_l, is defined as a vector of *k* single-locus genotypes that an individual carries. We denote by $S_l = \{i \mid X_i = G_l\}$ the group of subjects carrying a multi-locus genotype, G_l, $1 \le l \le L$; and $m_l = \mid S_l \mid$ the number of subjects in S_l.

For each single phenotype y_t, $1 \le t \le T$, we first choose a kernel function as $\varphi(y_{i,t}, y_{j,t}) = y_{i,t} - y_{j,t}$, and then define a general *L*-group U-Statistic,

$$U^{(t)} = \frac{\sum_{1 \le l < l' \le L} \omega_{l,l'} U^{(t)}_{l,l'}}{\sum_{1 \le l < l' \le L} \omega_{l,l'}} , \qquad \text{Eq. } (\mathbf{1})$$

where $U^{(t)}_{l,l'} = \sum_{i \in S_l, j \in S_{l'}} \varphi(y_{i,t}, y_{j,t})$ is a two-group U-statistic defined for groups S_l and $S_{l'}$, and $\omega_{l,l'} = \sqrt{(m_l + m_{l'})} / (m_l m_{l'})$ is a weight parameter to account for the number of subjects in various genotype groups. Given the U-Statistic of each phenotype defined in Eq. (1), a multivariate U-Statistic for *T* phenotypes, $Y = (y_1 ...... y_T)$, can be formed as $U = (U^{(1)}, U^{(2)} ...... U^{(T)})$. Under the null hypothesis of no association, it follows asymptotically a multivariate normal distribution, $N(\mathbf{0}, \Sigma)$. The test statistic to evaluate the joint association of *k* SNPs and *T* phenotypes is thus defined as:

$$\Delta = U \Sigma^{-1} U' , \qquad \text{Eq. } (\mathbf{3})$$

which follows a Chi-square distribution with *T* degrees of freedom, $\chi^2(T)$. In practice, the sample covariance matrix Σ is used in Eq. (**3**), which is detailed in the Appendix.

2.2. Forward Section of SNPs

While dealing with a large number of SNPs, it is likely that a significant proportion of SNPs are not disease-related. In this article, we follow the same strategy used in the forward U-test to select *k* disease-associated SNPs from total *K* genotyped SNPs, and use them to build the above test statistic Δ. This selection process starts with a single SNP. In the first step, each SNP *j* can partition the subjects into two genotype groups in three possible ways: $\{S_1 = \{i \mid x_{i,j} = \text{AA}\}, S_2 = \{i \mid x_{i,j} = \text{Aa/aa}\}\}$; $\{S_1 = \{i \mid x_{i,j} = \text{Aa}\}, S_2 = \{i \mid x_{i,j} = \text{AA/aa}\}\}$; and $\{S_1 = \{i \mid x_{i,j} = \text{aa}\}, S_2 = \{i \mid x_{i,j} = \text{AA/Aa}\}\}$.

We scanned each SNP, and select a single SNP and a partition with the maximum test statistic Δ. We denote the corresponding partitioning strategy in the first step as $\{S_1^{(1)}, S_2^{(1)}\}$. In the second step, a second SNP j' is selected and further partition all subjects into four groups, as $\{S_1^{(2)} = S_1^{(1)} \cap S_1^{j'}$, $S_2^{(2)} = S_1^{(1)} \cap S_2^{j'}$, $S_3^{(2)} = S_2^{(1)} \cap S_1^{j'}$, $S_4^{(2)} = S_2^{(1)} \cap S_2^{j'}\}$. Again, the SNP and partitioning strategy with the largest test statistic Δ is selected. It should be noted that under the null hypothesis of no association, the four groups of subjects (i.e. $S_1^{(2)} \ldots. S_4^{(2)}$) have the same phenotypic values. Upon the rejection of the null hypothesis, the phenotypes are not all equal among four groups of subjects, without assuming an additive effect from SNPs. Therefore, the proposed method tests for association while allowing for statistical interactions, [25] and SNPs with non-linear interaction effects can be detected. By repeating the selection process, SNPs is selected forwardly to partition subjects into multi-locus genotype groups. To avoid the issue of overfitting, a 10-fold cross-validation procedure is adopted to determine the most parsimonious model. Assuming the forward selection is stopped at step *s*, we then have the final model with *s* SNPs, which comprise *L* multi-locus genotype groups, $\{S_1^{(s)}, S_2^{(s)} \ldots\ldots, S_L^{(s)}\}$.

2.3. Backward Section of Phenotypes

When dealing with multiple phenotypes, it is also likely that a subset of phenotypes has no genetic relevance. Because the number of phenotypes is generally small, we propose to use a backward selection strategy to filter out phenotypes that are not genetically related. The selection process starts with all *T* available phenotypes. In the first step, multi-locus genotype groups can be formed by using the forward selection process described in Section 2.2, with a corresponding test statistic Δ_T. In the second step, by removing one phenotype, $y_t; 1 \le t \le T$, at a time, *T* possible phenotype subsets can be formed, each with *T*-1 phenotypes. For each subset of phenotypes, multi-locus genotype groups can be formed by using the forward selection, with a corresponding test statistic $\Delta_{T-1}^{(t)}, 1 \le t \le T$. The smallest test statistic obtained from *T* possible phenotype subsets, $\Delta_{T-1}^{(t_0)}$, will then be compared to that of *T* phenotypes, Δ_T by their corresponding p-values. The genotype-phenotype association can be assessed by $p_T = P(\chi^2(T) \ge \Delta_T)$ and $p_{T-1}^{(t_0)} = P(\chi^2(T-1) \ge \Delta_{T-1}^{(t_0)})$ for *T* phenotypes and *T*-1 phenotypes, respectively. We remove a phenotype, y_{t_0}, if $\Delta_{T-1}^{(t_0)}$ leads to a more significant association than Δ_T, i.e. $p_T \ge p_{T-1}^{(t_0)}$. The backward selection of phenotypes and forward selection of SNPs are conducted iteratively until no phenotypes can be removed to improve the significance of the association.

2.4. Test of Significance

Because the proposed method conducts model selection by maximizing the test statistic, the asymptotic test is no longer valid [26-28]. To examine the overall significance of the association, a permutation test is then conducted by randomly shuffling the phenotypes and then applying the forward selection of SNPs and backward selection of phenotypes as described above. Based on the permutation distribution of the test statistic Δ, an empirical *P*-value, which takes model selection into account, can be attained. In a replication study when the multi-locus genotype combinations and the subset of phenotypes are pre-determined from an initial study, the overall significance of the association can be obtained from a Chi-square distribution.

3. RESULTS

3.1. Simulation Studies

Simulation Settings

We conducted simulation studies to evaluate the proposed method, and compared it to the forward U-test, which analyzes one phenotype at a time. In each replicate, we simulated 1,000 subjects, each genotyped with 10 SNPs. The genotypes were simulated by assuming a minor allele frequency of 0.3 and Hardy Weinberg Equilibrium (HWE). Each simulation scenario was repeated for 1,000 times to evaluate the type I error rates and statistical power of two methods. For simplicity, we assume an additive model for *k*th SNP (i.e., $x_k = 0$ for AA, $x_k = 1$ for Aa, and $x_k = 2$ for aa). We first evaluated the type I error rate of the proposed method by simulating the phenotypes independently from the genotypes, assuming each phenotype follows a standard normal distribution. The type I error rates were evaluated for a varying number of phenotypes (i.e. from 1 to 5). To evaluate statistical power, phenotypes were simulated according to various disease scenarios described below.

Simulation I: Varying Number of Shared SNPs

In the first simulation, we considered two phenotypes, each influenced by 4 SNPs with an additive effect. We evaluated the performance of the proposed method by varying the number of shared SNPs that were associated with both phenotypes (i.e. from 0 to 4). The two phenotypes were thus simulated by:

$$\begin{cases} y_1 = \sum_{i=0}^{q} \alpha_i x_i + \sum_{j=0}^{4-q} \beta_j x_{q+j} + \varepsilon_1 \\ y_2 = \sum_{i=0}^{q} \alpha_i x_i + \sum_{k=0}^{4-q} \gamma_k x_{4+k} + \varepsilon_2 \end{cases};$$

$$\alpha_0 = \beta_0 = \gamma_0 = 0; \ q \in \{0,1,2,3,4\};$$

where the first q SNPs were shared SNPs that were associated with both phenotypes; and the remaining (4-q) SNPs were unique SNPs that were associated with one of the phenotypes. In such a disease scenario, we expected that the shared genetic components of two phenotypes would increase as the number of shared SNPs increased.

Simulation II: Varying Effect Size

In the second simulation, we also considered two phenotypes, each influenced by 2 SNPs with an additive effect. We further assumed two phenotypes shared one causal SNP. The phenotypes were thus simulated by:

$$\begin{cases} y_1 = \alpha x_1 + \beta x_2 + \varepsilon_1 \\ y_2 = \alpha x_1 + \gamma x_3 + \varepsilon_2 \end{cases};$$

where x_1 was the shared SNP that influenced both phenotypes, and x_2 and x_3 are unique SNPs that only influenced one of the phenotypes. We evaluated the performance of the proposed method by varying the relative contribution between the shared SNP and the unique SNPs (i.e. α/β and α/γ).

Simulation III: Varying Patterns of Interaction Effects

In the third simulation, we considered two phenotypes, influenced by 2 shared SNPs, but through various modes of inheritance that may or may not involve interactions. Each phenotype was simulated from three possible disease models, including an additive effect model, a multiplicative effect model, and a threshold effect model. The first model did not have an interaction effect, while the other two models had an interaction effect. The details of the simulation were described below:

a. Both phenotypes were simulated through an additive effect model,

$$\begin{cases} y_1 = \beta_1 x_1 + \beta_2 x_2 + \varepsilon_1 \\ y_2 = \beta_1 x_1 + \beta_2 x_2 + \varepsilon_2 \end{cases};$$

b. One phenotype was simulated through an additive effect model, while the other phenotype was simulated through a multiplicative effect model, which assumed an interaction effect on a multiplicative scale.

$$\begin{cases} y_1 = \beta_1 x_1 + \beta_2 x_2 + \varepsilon_1 \\ y_2 = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon_2 \end{cases};$$

c. One phenotype was simulated through an additive effect model, while the other phenotype was simulated through a threshold effect model, which assumed an interaction effect in the presence of minor alleles at both SNPs,

$$\begin{cases} y_1 = \beta_1 x_1 + \beta_2 x_2 + \varepsilon_1 \\ y_2 = \beta \times I(x_1 > 0) I(x_2 > 0) + \varepsilon_2 \end{cases};$$

d. One phenotype was simulated through a multiplicative effect model, while the other phenotype was simulated through a threshold effect model,

$$\begin{cases} y_1 = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon_1 \\ y_2 = \beta \times I(x_1 > 0) I(x_2 > 0) + \varepsilon_2 \end{cases}.$$

Simulation IV: Varying Number of Phenotypic Traits

In the fourth simulation, we considered a varying number of phenotypes (i.e. from 3 to 5). We further assumed only 2 phenotypes were genetically related, so that the number of noise phenotypes varied from 1 to 3. The first two phenotypes were simulated through the disease models discussed in Simulation III, while the remaining phenotypes were simulated independently from the genotypes, assuming a standard normal distribution.

3.2. Simulation Results

Type I Error

The results of type I error for the stepwise U-test are summarized in Table **1**. The results have shown that type I error of the new method remained well controlled at the level of 0.05 for different numbers of phenotypes.

Simulation I: Varying Number of Shared SNPs

To evaluate the statistical power, we conducted 1,000 permutation replicates for each simulation scenario. The power was defined as the probability of the observed test statistic exceeding the 95 percentile of the empirical permutation distribution. We also used sensitivity and specificity to measure the accuracy of SNP selection. In particular, *Sensitivity A* was defined as the probability to select a causal SNP that influenced only one of the phenotypes; *Sensitivity B* was defined as the probability to select a causal SNP that influ-

Table 1. Type I error rates of the stepwise U-test for different numbers of phenotypes.

Number of Phenotypes	1 Phenotypes	2 Phenotypes	3 Phenotypes	4 Phenotypes	5 Phenotypes
Type I error	0.042	0.053	0.045	0.051	0.050

enced both phenotypes; and *Specificity* was defined as 1 - the probability to select a SNP that influenced none of the phenotypes. The definition of these measurements remained same for all simulation scenarios.

The results of Simulation I are summarized in Table **2**. The results showed that the power of single-phenotype analysis remained stable around 0.50 (i.e. between 0.481 and 0.530). When two phenotypes shared no causal SNPs (i.e. q=0), the power of the multi-phenotype analysis (i.e. 0.544) was comparable to that of the single-phenotype analysis. However, the power of multi-phenotype analysis increased as the number of shared SNPs increased. When all causal SNPs were shared SNPs (i.e. q=4), the power of the multi-phenotype analysis (i.e. 0.903) was substantially higher than that of single-phenotype analysis. In terms of SNP selection, multi-phenotype showed an improved ability to select shared SNPs than single-phenotype analysis (i.e. sensitivity B), but reduced probability to select unique SNPs (i.e. sensitivity A). In terms of specificity, single-phenotype analysis and multi-phenotype analysis had comparable performance (i.e. around 95%).

Simulation II: Varying Effect Size

The results of Simulation II are summarized in Table **3**. When the effect sizes of causal SNPs increased, the statistical power of both multi-phenotype analysis and single-phenotype analysis increased. Furthermore, when the effect sizes of shared SNPs or unique SNPs increased, the power of single-phenotype analysis increased on a similar level. Nevertheless, the power of multi-phenotype analysis increased substantially when the effect size of shared SNP increased.

In terms of SNP selection, SNPs with larger effect sizes were more likely to be selected from either single-phenotype or multi-phenotype analysis. Multi-phenotype analysis may increase the probability to select a shared SNP (i.e. sensitivity A), but reduce the probability to select a unique SNP (i.e. sensitivity B). The specificity remained at a high level for both single-phenotype and multi-phenotype analyses (i.e. over 90%).

Simulation III: Varying Underlying Disease Models

The simulation results are summarized in Table **4**. The results showed that both single-phenotype and multi-phenotype analysis were able to detect the joint association when there was an interaction effect between SNPs. Furthermore, multi-phenotype analysis attained increased power over single-phenotype analysis. The power improvement was achieved with/without the interaction effect. In terms of SNP selection, multi-phenotype analysis had improved sensitivity and specificity over single-phenotype analysis for all scenarios.

Table 2. Power comparison between single-phenotype analyses and multi-phenotype analyses when the number of shared SNPs varies

Disease Model		Single-Phen[4]		Multi-Pheno[5]
		y_1	y_2	(y_1, y_2)
$\begin{cases} y_1 = 0.15x_1 + 0.15x_2 + 0.15x_3 + 0.15x_4 + \varepsilon_1 \\ y_2 = 0.15x_5 + 0.15x_6 + 0.15x_7 + 0.15x_8 + \varepsilon_2 \end{cases}$	Power	0.509	0.503	0.544
	Sensitivity A[1]	0.481	0.474	0.241
	Sensitivity B[2]	--	--	--
	Specificity[3]	0.960	0.956	0.959
$\begin{cases} y_1 = 0.15x_1 + 0.15x_2 + 0.15x_3 + 0.15x_4 + \varepsilon_1 \\ y_2 = 0.15x_1 + 0.15x_5 + 0.15x_6 + 0.15x_7 + \varepsilon_2 \end{cases}$	Power	0.489	0.509	0.626
	Sensitivity A	0.471	0.471	0.216
	Sensitivity B	0.486	0.476	0.628
	Specificity	0.957	0.962	0.967
$\begin{cases} y_1 = 0.15x_1 + 0.15x_2 + 0.15x_3 + 0.15x_4 + \varepsilon_1 \\ y_2 = 0.15x_1 + 0.15x_2 + 0.15x_5 + 0.15x_6 + \varepsilon_2 \end{cases}$	Power	0.514	0.513	0.769
	Sensitivity A	0.489	0.473	0.194
	Sensitivity B	0.462	0.479	0.666
	Specificity	0.961	0.961	0.973
$\begin{cases} y_1 = 0.15x_1 + 0.15x_2 + 0.15x_3 + 0.15x_4 + \varepsilon_1 \\ y_2 = 0.15x_1 + 0.15x_2 + 0.15x_3 + 0.15x_5 + \varepsilon_2 \end{cases}$	Power	0.527	0.530	0.880
	Sensitivity A	0.463	0.491	0.152
	Sensitivity B	0.476	0.486	0.675
	Specificity	0.959	0.962	0.974
$\begin{cases} y_1 = 0.15x_1 + 0.15x_2 + 0.15x_3 + 0.15x_4 + \varepsilon_1 \\ y_2 = 0.15x_1 + 0.15x_2 + 0.15x_3 + 0.15x_4 + \varepsilon_2 \end{cases}$	Power	0.491	0.481	0.903
	Sensitivity A	--	--	--
	Sensitivity B	0.481	0.466	0.677
	Specificity	0.960	0.955	0.973

[1] Sensitivity A: the probability of selecting a causal SNP that influences only one phenotype

[2] Sensitivity B: the probability of selecting a causal SNP that influences both phenotypes

[3] Specificity : the probability of selecting a SNP that influences none of the phenotypes

[4] single-phenotype analyses are conducted by using forward U-test

[5] multi-phenotype analyses are conducted by using stepwise U-test

Table 3. Power comparison between single-phenotype analysis and multi-phenotype analysis when the effect sizes vary

Disease Model		**Single-Phen**[1]		**Multi-Phen**[2]
		y_1	y_2	(y_1, y_2)
$\begin{cases} y_1 = 0.1x_1 + 0.1x_2 + \varepsilon_1 \\ y_2 = 0.1x_1 + 0.1x_3 + \varepsilon_2 \end{cases}$	Power Sensitivity A Sensitivity B Specificity	0.191 0.439 0.456 0.903	0.178 0.473 0.462 0.900	0.353 0.260 0.600 0.913
$\begin{cases} y_1 = 0.1x_1 + 0.2x_2 + \varepsilon_1 \\ y_2 = 0.1x_1 + 0.2x_3 + \varepsilon_2 \end{cases}$	Power Sensitivity A Sensitivity B Specificity	0.320 0.944 0.253 0.954	0.340 0.929 0.238 0.952	0.449 0.505 0.281 0.959
$\begin{cases} y_1 = 0.2x_1 + 0.1x_2 + \varepsilon_1 \\ y_2 = 0.2x_1 + 0.1x_3 + \varepsilon_2 \end{cases}$	Power Sensitivity A Sensitivity B Specificity	0.330 0.246 0.940 0.950	0.335 0.233 0.941 0.962	0.823 0.087 0.996 0.971
$\begin{cases} y_1 = 0.2x_1 + 0.2x_2 + \varepsilon_1 \\ y_2 = 0.2x_1 + 0.2x_3 + \varepsilon_2 \end{cases}$	Power Sensitivity A Sensitivity B Specificity	0.645 0.845 0.832 0.960	0.650 0.823 0.846 0.954	0.927 0.371 0.916 0.971

[1] single-phenotype analysis is conducted by using forward U-test

[2] multi-phenotype analysis is conducted by using stepwise U-test

Table 4. Power comparison between single-phenotype analysis and multi-phenotype analysis by varying underlying disease models

Disease Model		**Single-Phen**[1]		**Multi-Phen**[2]
		y_1	y_2	(y_1, y_2)
$\begin{cases} y_1 = 0.1x_1 + 0.1x_2 + \varepsilon_1 \\ y_2 = 0.1x_1 + 0.1x_2 + \varepsilon_2 \end{cases}$	Power Sensitivity Specificity	0.167 0.453 0.900	0.181 0.469 0.904	0.404 0.582 0.933
$\begin{cases} y_1 = 0.1x_1 + 0.1x_2 + \varepsilon_1 \\ y_2 = 0.2I(x_1 > 0)I(x_2 > 0) + \varepsilon_2 \end{cases}$	Power Sensitivity Specificity	0.165 0.457 0.904	0.124 0.378 0.883	0.343 0.526 0.916
$\begin{cases} y_1 = 0.1x_1 + 0.1x_2 + \varepsilon_1 \\ y_2 = 0.1x_1 + 0.1x_2 + 0.05x_1x_2 + \varepsilon_2 \end{cases}$	Power Sensitivity Specificity	0.162 0.447 0.900	0.280 0.567 0.930	0.512 0.660 0.945
$\begin{cases} y_1 = 0.1x_1 + 0.1x_2 + 0.05x_1x_2 + \varepsilon_1 \\ y_2 = 0.3 \times I(x_1 > 0)I(x_2 > 0) + \varepsilon_2 \end{cases}$	Power Sensitivity Specificity	0.314 0.582 0.930	0.334 0.600 0.934	0.722 0.744 0.953

[1] single-phenotype analysis is conducted by using forward U-test

[2] multi-phenotype analysis is conducted by using stepwise U-test

Simulation IV: Varying Number of Phenotypes

The simulation results are summarized in Table **5**. The results showed that the power decreased slightly as the number of noise phenotypes increased. In terms of SNP selection, both sensitivity and specificity decreased when the number of noise phenotypes increased.

In summary, our simulations have shown that: 1) Compared to the analysis of single phenotype with forward U-test, the analysis of multiple phenotypes with stepwise U-test has increased power to detect the association, especially when the phenotypes share relatively large genetic causes (e.g. more shared SNPs, larger effect size of shared SNPs). 2) Stepwise U-test has an increased the probability to detect shared SNPs, but a reduced probability to detect SNPs that are only causal to a particular phenotype. 3) Similar to forward U-test, stepwise U-test is able to detect the joint association when there are genetic interactions between genetic

Table 5. Performance of multi-phenotype analysis with varying number of noise phenotypes

Disease Model		2 Pheno	+1 noise	+2 noise	+3 noise
$\begin{cases} y_1 = 0.2x_1 + 0.2x_3 + \varepsilon_1 \\ y_2 = 0.2x_2 + 0.2x_3 + \varepsilon_2 \end{cases}$	Power	0.927	0.922	0.906	0.852
	Sensitivity A	0.371	0.348	0.344	0.339
	Sensitivity B	0.916	0.896	0.892	0.881
	Specificity	0.971	0.963	0.956	0.950
$\begin{cases} y_1 = 0.1x_1 + 0.1x_2 + 0.05x_1x_2 + \varepsilon_1 \\ y_2 = 0.3 \times I(x_1 > 0)I(x_2 > 0) + \varepsilon_2 \end{cases}$	Power	0.722	0.716	0.653	0.570
	Sensitivity	0.744	0.676	0.663	0.638
	Specificity	0.953	0.965	0.951	0.938

variants. 4) The performance of stepwise U-test remains robust in the presence of noise phenotypes.

3.3. Application to a Nicotine Dependence (ND) Dataset

We illustrated the proposed stepwise U-test with an application to a dataset from the Study of Addiction: Genetics and Environment (SAGE). The SAGE study is part of the Gene Environment Association Studies initiative (GENEVA) funded by the National Human Genome Research Institute. The SAGE samples were selected from three large complementary datasets: the Family Study of Cocaine Dependence (FSCD), the Collaborative Study on the Genetics of Alcoholism (COGA), and the Collaborative Genetic Study of Nicotine Dependence (COGEND) [29]. All samples in SAGE were unrelated and have quantitative measurements of various phenotypes for additions, such as alcohol, nicotine, marijuana, cocaine, opiates and other drugs. In this article, we focused on three ND-related phenotypes, including participant's lifetime score on Fagerström Test for Nicotine Dependence (ftnd_total), number of cigarettes smoked per day (ftnd_4), and number of nicotine symptoms endorsed (nic_sx_tot). We evaluated the joint association between three phenotypes and 155 SNPs that were reported for their potential association with ND. Because the SAGE study only had the genotypes of 128 SNPs, we further imputed the genotype of the other 27 SNPs by using PLINK [30]. Our study population was mainly biracial, and we used HapMap phase III founders of CEU (Utah residence with Northern and Western European ancestry) and ASW (African ancestry in Southwest USA) as the reference panels for the Caucasian and African American subjects respectively [31].

We applied stepwise U-test to samples of COGEND for an initial association analysis and to samples of FSCD and COGA for the replication analysis. The results are summarized in Table **6**. Based on the initial dataset COGEND, the analysis identified two SNPs, rs10508649 and rs2491397, joint associated with three ND-related phenotypes, with a nominal *P*-value of 3.79e-13. By using permutation, the empirical p-value of the association reached the significance level of 0.001. This association remained to be significant in both FSCD (*P*-value=2.37e-05) and COGA (*P*-value=7.46e-05).

For comparison purposes, we also conducted single-phenotype analyses by using forward U-test. The findings of single-phenotype analyses varied among three phenotypes. Based on the initial dataset of COGEND, 1) the analysis of the lifetime FTND score (ftnd_total) identified the same two SNPs with the multi-phenotype analyses; 2) the analysis of the number of cigarettes smoked *per* day (ftnd_4) revealed a different SNP, rs2036527; 3) the analysis of the number of nicotine symptoms endorsed (nic_sx_tot) found two SNPs, rs10508649 and rs7517376, one of which overlapped with the SNPs identified from the multi-phenotype analyses. All of the findings from single-phenotype analyses showed significant associations in the initial data COGEND. However, these associations could not be replicated in either FSCD or COGA. This result indicated that the proposed multi-phenotype strategy might improve the testing power and obtain more robust findings over its single-phenotype alternative.

4. DISCUSSION

Complex diseases are thought to be influenced by the interplay of hundreds or even thousands genetic variants through complex mechanisms [32]. Multi-locus methods, taking genetic interactions into account, could have improved power to detect disease-susceptibility genetic variants. Furthermore, complex phenotypes, such as nicotine dependence, are commonly assessed by multiple measures that are complementary to each other [33-37]. For example, the two gold-standard measures of nicotine dependence, the FTND score and the Diagnostic and Statistical Manual of Mental Disorders (DSM), were found to have a relatively low concordance with a Kappa estimate of 0.2. [35, 38] It was suggested that the FTND and DSM measurements emphasis on physical symptoms and psychiatric symptoms, respectively, each of which reflects a unique aspect of ND development. Other studies have also pointed out that ND can be assessed through various aspects, including physical, behavioral and psychological components [36]. While it remains challenging to define a single comprehensive measurement for better characterizing complex phenotypes, such as ND, new statistical methods can be used to facilitate the genetic discovery process by taking advantage of currently available multiple phenotypes in the analysis.

In this article, we proposed a stepwise U-test for testing the joint association between multiple loci and multiple phenotypes. Similar to the forward U-test developed for single-phenotype analyses, the proposed method is entirely nonparametric, which makes no assumption of the phenotype distribution and the underlying disease mechanisms (i.e.

Table 6. Summary of multi-phenotype analysis and single-phenotype analysis of three independent datasets, COGEND, FSCD and COGA.

Phenotype	SNP	Allele	Gene	Grouping	*P*-values
Multiple Phenotype Analyses					
3 Phenotypes	rs10508649 rs2491397	C/T C/T	*PIP*4*K*2*A* *GABBR*2	TT or CC/CT CC or CT/TT	COGEND: 3.79e-13 FSCD: 2.37e-05 COGA: 7.46e-05
Single Phenotype Analyses					
FTND_4	rs2036527	A/G	*CHRNA*5	AA or AG/GG	COGEND: 3.06e-05 FSCD: 0.180 COGA: 0.219
FTND_total	rs10508649 rs2491397	C/T C/T	*PIP4K2A* *GABBR*2	TT or CC/CT CC or CT/TT	COGEND: 1.39e-07 FSCD: 0.228 COGA: 0.066
Nic_sx_tot	rs10508649 rs7517376	C/T A/G	*PIP*4*K*2*A* *FMO*1	TT or CC/CT AA or AG/GG	COGEND: 4.92e-07 FSCD: 0.056 COGA: 0.526

modes of inheritance). We conducted simulation studies to compare the performance of two testing strategies: single-phenotype analysis and multi-phenotype analysis. Our simulation results demonstrated that multi-phenotype analysis could have better performance than single-phenotype analysis, especially when phenotypes of interest have similar underlying genetic etiologies (e.g., share part of causal genetic variants). The better performance of multi-phenotype analysis can be explained by its capacity of capturing collective effect of genetic variants over all relevant phenotypes. When there are a significant number of shared SNPs contributed to these phenotypes, multi-phenotype analysis is expect to outperform single-phenotype analysis.

In the article, we have focused on the association analysis of multiple quantitative phenotypes, by using a kernel function to measure the phenotype difference between two subjects. For dichotomous phenotypes, extension can be made by using a different kernel function [39]:

$$\varphi[LR(G_i), LR(G_j)] = \begin{cases} 1 & \text{if } LR(G_j) < LR(G_i) \\ 0.5 & \text{if } LR(G_j) = LR(G_i) \\ 0 & \text{if } LR(G_j) > LR(G_i) \end{cases};$$

where $LR(G) = \dfrac{P(G \mid y=1)}{P(G \mid y=0)}$ is the likelihood ratio of a particular genotype group.

Similar to forward U-test for the analysis of single phenotype, stepwise U-test also adopted a forward search strategy to select disease-associated SNPs. Therefore, it is computationally feasible to apply stepwise U-test to a relatively large number of SNPs. The computational time will depend on various factors, such as the number of variants, the number of phenotypes and sample size. Under our simulation setting with 2 phenotypes, 10 SNPs and 1,000 samples, it took an average computation time of 128.4 (SD=71.6) seconds to run each replicate on a desktop with a single core of 2.90 GHz and 8 GB RAM.

In the real data application, we identified and replicated the joint association of two SNPs, rs10508649 and rs2491397, with three ND phenotypes based on three independent datasets. The two SNPs are located in two genes, *PIP*4*K*2*A* and *GABBR*2, respectively. Gene *GABAB*2, known as Gamma-aminobutyric acid (GABA) B receptor 2, is a G-protein coupled receptor subunit that mediates inhibitory neurotransmitter in the central nervous system [40]. SNP rs2491397 was reported to be associated with the development of ND through haplotypes in the *GABAR*2 gene, [41] which was found to be associated with a number of measurements of ND, including the smoking quantity (SQ), the heaviness of smoking index (HSI), and the FTND score [42-44]. *GABAB*2 was also reported to be interacting with other genes, such as *GABAB*1, for a joint association with ND [45]. Moreover, *PIP*4*K*2*A* was found to be associated with other psychiatric disorders, such as schizophrenia [46-49]. SNP rs10508649 is located within *PIP*4*K*2*A*, and was found to be associated with ND outcome measured by FTND score [50]. In our study, the results also indicated that this SNP was potentially associated with other ND measurements, such as the number of symptoms endorsed. While it is biologically plausible that these two identified SNPs may be involved in a number of manifestations of ND, further studies are still needed to replicate the findings and investigate their effects on ND development.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

This research was supported, in part, by the NIH National Institute on Drug Abuse under Award Number K01DA

033346, the National Institute of Dental & Craniofacial Research under Award Number R03DE022379, the Translational Research Institute through the NIH National Center for Research Resources and the National Center for Advancing Translational Sciences under Award Number UL1TR000039 and KL2TR000063, the University of Arkansas for Medical Sciences College of Medicine Children's University Medical Group Fund Grant Program, the Arkansas Children's Hospital Research Institute and the Arkansas Bioscience Institute. The content is solely the responsibility of authors and does not necessarily represent the official views of NIH.

Funding support for the Study of Addiction: Genetics and Environment (SAGE) was provided through the NIH Genes, Environment and Health Initiative [GEI] (U01 HG004422). The SAGE datasets used for the analyses were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/projects/gap/cgibin/study.cgi?study_id=phs000092.v1.p1 through dbGaP accession number phs000092.v1.p1.

APPENDIX

EMPIRICAL ESTIMATION OF THE VARIANCE-COVARIANCE MATRIX OF U

Suppose we have a study population of *N* subjects, each measured with *T* phenotypes, $Y=(y_1......y_T)$. We assume the subjects are independent, and denote $Var(y_t)=\sigma_t^2$, $Cov(y_{t_1},y_{t_2})=\sigma_{t_1,t_2}$, where $1\le t,t_1,t_2\le T$. The proposed multivariate U-Statistic has a form of $U=(U^{(1)},......,U^{(T)})$. For simplicity, we denote $U^{(t)}=\sum_{1\le l<l'\le L}\alpha_{l,l'}^{(t)}U_{l,l'}^{(t)}$, where $U_{l,l'}^{(t)}=\sum_{i\in S_l,\,j\in S_{l'}}\varphi(y_{i,t},y_{j,t})$, $1\le t\le T$.

a) Variance of univariate U-Statistic for each phenotype

For any *t*, $1\le t\le T$, the derivation of $Var(U^{(t)})$ was detailed elsewhere [23]. Following the same notation, we have

$$Var(U^{(t)})=Var(\sum_{1\le l<l'\le L}\alpha_{l,l'}^{(t)}U_{l,l'}^{(t)})=\sum_{1\le l<l'\le L}\alpha_{l,l'}^{(t)}Var(U_{l,l'}^{(t)})+\sum_{\substack{1\le l_1<l_1'\le L\\1\le l_2<l_2'\le L\\(l_1,l_1')\ne(l_2,l_2')}}\alpha_{l_1,l_1'}^{(t)}\alpha_{l_2,l_2'}^{(t)}Cov(U_{l_1,l_1'}^{(t)},U_{l_2,l_2'}^{(t)})\,.$$

For all $1\le l<l'\le L$, we have:

$$\begin{aligned}Var(U_{l,l'}^{(t)})&=Var(\sum_{i\in S_l,\,j\in S_{l'}}\varphi(y_{i,t},y_{j,t}))\\&=Var(\sum_{i\in S_l,\,j\in S_{l'}}(y_{i,t}-y_{j,t}))\\&=Var(m_{l'}\sum_{i\in S_l}y_{i,t}-m_l\sum_{j\in S_{l'}}y_{j,t})\\&=m_{l'}^2\sum_{i\in S_l}Var(y_{i,t})+m_l^2\sum_{j\in S_{l'}}Var(y_{j,t})\\&=(m_{l'}^2m_l+m_l^2m_{l'})\ \sigma_t^2\end{aligned}\,.$$

The covariance, $Cov(U_{l_1,l_1'}^{(t)},U_{l_2,l_2'}^{(t)})$, can be estimated according to different scenarios:

1) when $l_1\ne l_1'\ne l_2\ne l_2'$,

$Cov(U_{l_1,l_1'}^{(t)},U_{l_2,l_2'}^{(t)})=0$;

2) when $l_1=l_2=l$

$$
\begin{aligned}
Cov(U^{(t)}_{l_1,l'_1}, U^{(t)}_{l_2,l'_2}) &= Cov(U^{(t)}_{l\ ,l'_1}, U^{(t)}_{l\ ,l'_2}) \\
&= Cov(\sum_{i\in S_l, j_1\in S_{l'_1}} \varphi(y_{i,t}, y_{j_1,t}), \sum_{i\in S_l, j_2\in S_{l'_2}} \varphi(y_{i,t}, y_{j_2,t})) \\
&= Cov(\sum_{i\in S_l, j_1\in S_{l'_1}} (y_{i,t} - y_{j_1,t}), \sum_{i\in S_l, j_2\in S_{l'_2}} (y_{i,t} - y_{j_2,t})) \\
&= Cov(m_{l'_1} \sum_{i\in S_l} y_{i,t}, m_{l'_2} \sum_{i\in S_l} y_{j_2,t}) \\
&= m_{l'_1} m_{l'_2} Var(\sum_{i\in S_l} y_{i,t}) \\
&= m_{l'_1} m_{l'_2} m_l\ \sigma_t^2
\end{aligned}
;
$$

3) when $l_1' = l_2' = l$,

$$
\begin{aligned}
Cov(U^{(t)}_{l_1,l'_1}, U^{(t)}_{l_2,l'_2}) &= Cov(U^{(t)}_{l_1,l}\ , U^{(t)}_{l_2,l}\ ) \\
&= Cov(\sum_{i_1\in S_{l_1}, j\in S_l} \varphi(y_{i_1,t}, y_{j,t}), \sum_{i_2\in S_{l_2}, j\in S_l} \varphi(y_{i_2,t}, y_{j,t})) \\
&= Cov(\sum_{i_1\in S_{l_1}, j\in S_l} (y_{i_1,t} - y_{j,t}), \sum_{i_2\in S_{l_2}, j\in S_l} (y_{i_2,t} - y_{j,t})) \\
&= Cov(m_{l_1} \sum_{j\in S_l} y_{j,t}, m_{l_2} \sum_{j\in S_l} y_{j,t}) \\
&= m_{l_1} m_{l_2} Var(\sum_{j\in S_l} y_{j,t}) \\
&= m_{l_1} m_{l_2} m_l\ \sigma_t^2
\end{aligned}
;
$$

4) when $l_1' = l_2 = l$ or $l_1 = l_2' = l$,

$$
\begin{aligned}
Cov(U^{(t)}_{l_1,l'_1}, U^{(t)}_{l_2,l'_2}) &= Cov(U^{(t)}_{l_1,l}\ , U^{(t)}_{l,l'_2}) \\
&= Cov(\sum_{i\in S_{l_1}, j\in S_l} \varphi(y_{i,t}, y_{j,t}), \sum_{j\in S_l, k\in S_{l'_2}} \varphi(y_{j,t}, y_{k,t})) \\
&= Cov(\sum_{i\in S_{l_1}, j\in S_l} (y_{i,t} - y_{j,t}), \sum_{j\in S_l, t\in S_{l'_2}} (y_{j,t} - y_{k,t})) \\
&= Cov(-m_{l'_1} \sum_{j\in S_l} y_{j,t}, m_{l'_2} \sum_{j\in S_l} y_{j,t}) \\
&= -m_{l'_1} m_{l'_2} Var(\sum_{j\in S_l} y_{j,t}) \\
&= -m_{l_1} m_{l'_2} m_l\ \sigma_t^2
\end{aligned}
.
$$

b) Covariance of the univariate U-Statistics of two phenotypes

For $1 \le t_1 < t_2 \le T$, we have:

$$
\begin{aligned}
Cov(U^{(t_1)}, U^{(t_2)}) &= Cov(\sum_{1\le l_1 < l'_1 \le L} \alpha^{(t_1)}_{l_1,l'_1}\ U^{(t_1)}_{l_1,l'_1}, \sum_{1\le l_2 < l'_2 \le L} \alpha^{(t_2)}_{l_2,l'_2}\ U^{(t_2)}_{l_2,l'_2}) \\
&= \sum_{1\le l_1 < l'_1 \le L} \sum_{1\le l_2 < l'_2 \le L} \alpha^{(t_1)}_{l_1,l'_1}\ \alpha^{(t_2)}_{l_2,l'_2} Cov(U^{(t_1)}_{l_1,l'_1}, U^{(t_2)}_{l_2,l'_2})
\end{aligned}
.
$$

Further, the co-variance can be estimated according to different scenarios:

1) when $l_1 = l_2 < l'_1 = l'_2$, we denote $l_1 = l_2 = l$ and $l'_1 = l'_2 = l'$,

$$\begin{aligned}
Cov(U^{(t_1)}_{l_1,l'_1}, U^{(t_2)}_{l_2,l'_2}) &= Cov(U^{(t_1)}_{l,l'}, U^{(t_2)}_{l,l'}) \\
&= Cov(\sum_{i \in S_l\ ,\, j \in S_{l'}} \varphi(y_{i,t_1}, y_{j,t_1}), \sum_{i \in S_l\ ,\, j \in S_{l'}} \varphi(y_{i,t_2}, y_{j,t_2})) \\
&= Cov(\sum_{i \in S_l\ ,\, j \in S_{l'}} (y_{i,t_1} - y_{j,t_1}), \sum_{i \in S_l\ ,\, j \in S_{l'}} (y_{i,t_2} - y_{j,t_2})) \\
&= Cov(\sum_{i \in S_l\ ,} y_{i,t_1}, \sum_{i \in S_l\ ,} y_{i,t_2}) + Cov(\sum_{j \in S_{l'}\ ,} y_{j,t_1}, \sum_{j \in S_{l'}\ ,} y_{j,t_2}) \\
&= m_l^2 \sigma_{t_1,t_2} + m_{l'}^2 \sigma_{t_1,t_2} \\
&= (m_l^2 + m_{l'}^2)\ \sigma_{t_1,t_2}
\end{aligned}\qquad ;$$

2) when $l_1 = l_2 < l'_1 < l'_2$ or $l_1 = l_2 < l'_2 < l'_1$, we denote $l_1 = l_2 = l$,

$$\begin{aligned}
Cov(U^{(t_1)}_{l_1,l'_1}, U^{(t_2)}_{l_2,l'_2}) &= Cov(U^{(t_1)}_{l,l'_1}, U^{(t_2)}_{l,l'_2}) \\
&= Cov(\sum_{i \in S_l\ ,\, j_1 \in S_{l'_1}} \varphi(y_{i,t_1}, y_{j,t_1}), \sum_{i \in S_l\ ,\, j_2 \in S_{l'_2}} \varphi(y_{i,t_2}, y_{j,t_2})) \\
&= Cov(\sum_{i \in S_l\ ,\, j_1 \in S_{l'_1}} (y_{i,t_1} - y_{j_1,t_1}), \sum_{i \in S_l\ ,\, j_2 \in S_{l'_2}} (y_{i,t_2} - y_{j_2,t_2})) \\
&= Cov(\sum_{i \in S_l\ ,} y_{i,t_1}, \sum_{i \in S_l\ ,} y_{i,t_2}) \\
&= m_l^2 \sigma_{t_1,t_2}
\end{aligned}\qquad ;$$

3) when $l_1 < l'_1 = l_2 < l'_2$ or $l_2 < l'_2 = l_1 < l'_1$, we denote $l'_1 = l_2 = l$,

$$\begin{aligned}
Cov(U^{(t_1)}_{l_1,l'_1}, U^{(t_2)}_{l_2,l'_2}) &= Cov(U^{(t_1)}_{l_1,l}, U^{(t_2)}_{l,l'_2}) \\
&= Cov(\sum_{i_1 \in S_{l_1},\, j \in S_l} \varphi(y_{i_1,t_1}, y_{j,t_1}), \sum_{j \in S_l\ ,\, i_2 \in S_{l'_2}} \varphi(y_{j,t_2}, y_{i_2,t_2})) \\
&= Cov(\sum_{i_1 \in S_{l_1},\, j \in S_l} (y_{i_1,t_1} - y_{j,t_1}), \sum_{j \in S_l\ ,\, i_2 \in S_{l'_2}} (y_{j,t_2} - y_{i_2,t_2})) \\
&= Cov(\sum_{j \in S_l\ ,} -y_{j,t_1}, \sum_{j \in S_l\ ,} y_{j,t_2}) \\
&= -m_l^2 \sigma_{t_1,t_2}
\end{aligned}\qquad ;$$

4) for others,

$$Cov(U^{(t_1)}_{l_1,l'_1}, U^{(t_2)}_{l_2,l'_2}) = 0 \;.$$

REFERENCES

[1] Hindorff, L. A.; Sethupathy, P.; Junkins, H. A.; Ramos, E. M.; Mehta, J. P.; Collins, F. S.; Manolio, T. A. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. U. S. A.,* **2009**, *106* (23), 9362-7.

[2] Welter, D.; MacArthur, J.; Morales, J.; Burdett, T.; Hall, P.; Junkins, H.; Klemm, A.; Flicek, P.; Manolio, T.; Hindorff, L.; Parkinson, H. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.,* **2014**, *42* (Database issue), D1001-6.

[3] Maher, B., Personal genomes: The case of the missing heritability. *Nature,* **2008**, *456* (7218), 18-21.

[4] Manolio, T. A.; Collins, F. S.; Cox, N. J.; Goldstein, D. B.; Hindorff, L. A.; Hunter, D. J.; McCarthy, M. I.; Ramos, E. M.; Cardon, L. R.; Chakravarti, A.; Cho, J. H.; Guttmacher, A. E.; Kong, A.; Kruglyak, L.; Mardis, E.; Rotimi, C. N.; Slatkin, M.; Valle, D.; Whittemore, A. S.; Boehnke, M.; Clark, A. G.; Eichler, E. E.; Gibson, G.; Haines, J. L.; Mackay, T. F.; McCarroll, S. A.; Visscher, P. M. Finding the missing heritability of complex diseases. *Nature,* **2009**, *461* (7265), 747-53.

[5] So, H. C.; Gui, A. H.; Cherny, S. S.; Sham, P. C. Evaluating the heritability explained by known susceptibility variants: a survey of ten complex diseases. *Genet. Epidemiol.,* **2011**, *35* (5), 310-7.

[6] Pendergrass, S. A.; Brown-Gentry, K.; Dudek, S. M.; Torstenson, E. S.; Ambite, J. L.; Avery, C. L.; Buyske, S.; Cai, C.; Fesinmeyer,

M. D.; Haiman, C.; Heiss, G.; Hindorff, L. A.; Hsu, C. N.; Jackson, R. D.; Kooperberg, C.; Le Marchand, L.; Lin, Y.; Matise, T. C.; Moreland, L.; Monroe, K.; Reiner, A. P.; Wallace, R.; Wilkens, L. R.; Crawford, D. C.; Ritchie, M. D. The use of phenome-wide association studies (PheWAS) for exploration of novel genotype-phenotype relationships and pleiotropy discovery. *Genet. Epidemiol.,,* **2011**, *35* (5), 410-22.

[7] Stearns, F. W. One hundred years of pleiotropy: a retrospective. *Genetic,* **2010**, *186* (3), 767-73.

[8] Ritchie, M. D.; Denny, J. C.; Crawford, D. C.; Ramirez, A. H.; Weiner, J. B.; Pulley, J. M.; Basford, M. A.; Brown-Gentry, K.; Balser, J. R.; Masys, D. R.; Haines, J. L.; Roden, D. M. Robust replication of genotype-phenotype associations across multiple diseases in an electronic medical record. *Am. J. Hum. Genet.,* **2010**, *86* (4), 560-72.

[9] Roden, D. M.; Pulley, J. M.; Basford, M. A.; Bernard, G. R.; Clayton, E. W.; Balser, J. R.; Masys, D. R. Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin. Pharmacol. Ther.,* **2008**, *84* (3), 362-9.

[10] Lange, C.; van Steen, K.; Andrew, T.; Lyon, H.; DeMeo, D. L.; Raby, B.; Murphy, A.; Silverman, E. K.; MacGregor, A.; Weiss, S. T.; Laird, N. M. A family-based association test for repeatedly measured quantitative traits adjusting for unknown environmental and/or polygenic effects. *Stat. Appl. Genet. Mol. Biol.,* **2004**, *3*, Article17.

[11] Klei, L.; Luca, D.; Devlin, B.; Roeder, K. Pleiotropy and principal components of heritability combine to increase power for association analysis. *Genet. Epidemiol.,* **2008**, *32* (1), 9-19.

[12] Lange, C.; Silverman, E. K.; Xu, X.; Weiss, S. T.; Laird, N. M. A multivariate family-based association test using generalized estimating equations: FBAT-GEE. *Biostatistics,* **2003**, *4* (2), 195-206.

[13] Zhang, H.; Liu, C. T.; Wang, X. An Association Test for Multiple Traits Based on the Generalized Kendall's Tau. *J. Am. Stat. Assoc.,* **2010**, *105* (490), 473-481.

[14] Nelson, M. R.; Kardia, S. L.; Ferrell, R. E.; Sing, C. F. A combinatorial partitioning method to identify multilocus genotypic partitions that predict quantitative trait variation. *Genome Res.,* **2001**, *11* (3), 458-70.

[15] Ritchie, M. D.; Hahn, L. W.; Roodi, N.; Bailey, L. R.; Dupont, W. D.; Parl, F. F.; Moore, J. H. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am. J. Hum. Genet.,* **2001**, *69* (1), 138-47.

[16] Culverhouse, R.; Klein, T.; Shannon, W. Detecting epistatic interactions contributing to quantitative traits. *Genet. Epidemiol.,* **2004**, *27* (2), 141-52.

[17] Sha, Q.; Zhu, X.; Zuo, Y.; Cooper, R.; Zhang, S. A combinatorial searching method for detecting a set of interacting loci associated with complex traits. *Ann. Hum. Genet.,* **2006**, *70* (Pt 5), 677-92.

[18] Lou, X. Y.; Chen, G. B.; Yan, L.; Ma, J. Z.; Zhu, J.; Elston, R. C.; Li, M. D., A generalized combinatorial approach for detecting gene-by-gene and gene-by-environment interactions with application to nicotine dependence. *Am. J. Hum. Genet.,* **2007**, *80* (6), 1125-37.

[19] Kwee, L. C.; Liu, D.; Lin, X.; Ghosh, D.; Epstein, M. P., A powerful and flexible multilocus association test for quantitative traits. *Am. J. Hum. Genet.,* **2008**, *82* (2), 386-97.

[20] Zhang, F.; Guo, X.; Deng, H. W., Multilocus association testing of quantitative traits based on partial least-squares analysis. *PLoS One,* **2011**, *6* (2), e16739.

[21] Schaid, D. J.; McDonnell, S. K.; Hebbring, S. J.; Cunningham, J. M.; Thibodeau, S. N., Nonparametric tests of association of multiple genes with human disease. *Am. J. Hum. Genet.,* **2005**, *76* (5), 780-93.

[22] Wei, Z.; Li, M.; Rebbeck, T.; Li, H., U-statistics-based tests for multiple genes in genetic association studies. *Ann. Hum. Genet.,* **2008**, *72* (Pt 6), 821-33.

[23] Li, M.; Ye, C.; Fu, W.; Elston, R. C.; Lu, Q., Detecting genetic interactions for quantitative traits with U-statistics. *Genet. Epidemiol.,* **2011**, *35* (6), 457-68.

[24] Lu, Q.; Wei, C.; Ye, C.; Li, M.; Elston, R. C., A likelihood ratio-based Mann-Whitney approach finds novel replicable joint gene action for type 2 diabetes. *Genet. Epidemiol.,* **2012**, *36* (6), 583-93.

[25] Cordell, H. J., Detecting gene-gene interactions that underlie human diseases. *Nat. Rev. Genet.,* **2009**, *10* (6), 392-404.

[26] Edgington, E. S., *Randomizatin tests. 3rd ed.* Dekker: New York., **1995.**

[27] Churchill, G. A.; Doerge, R. W., Empirical threshold values for quantitative trait mapping. *Genetics,* **1994**, *138* (3), 963-71.

[28] Zhao, J. H.; Curtis, D.; Sham, P. C., Model-free analysis and permutation tests for allelic associations. *Hum. Hered.,* **2000**, *50* (2), 133-9.

[29] Bierut, L. J.; Agrawal, A.; Bucholz, K. K.; Doheny, K. F.; Laurie, C.; Pugh, E.; Fisher, S.; Fox, L.; Howells, W.; Bertelsen, S.; Hinrichs, A. L.; Almasy, L.; Breslau, N.; Culverhouse, R. C.; Dick, D. M.; Edenberg, H. J.; Foroud, T.; Grucza, R. A.; Hatsukami, D.; Hesselbrock, V.; Johnson, E. O.; Kramer, J.; Krueger, R. F.; Kuperman, S.; Lynskey, M.; Mann, K.; Neuman, R. J.; Nothen, M. M.; Nurnberger, J. I., Jr.; Porjesz, B.; Ridinger, M.; Saccone, N. L.; Saccone, S. F.; Schuckit, M. A.; Tischfield, J. A.; Wang, J. C.; Rietschel, M.; Goate, A. M.; Rice, J. P., A genome-wide association study of alcohol dependence. *Proc. Natl. Acad. Sci. U. S. A.,* **2010**, *107* (11), 5082-7.

[30] Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M. A.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P. I.; Daly, M. J.; Sham, P. C., PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.,* **2007**, *81* (3), 559-75.

[31] Altshuler, D. M.; Gibbs, R. A.; Peltonen, L.; Dermitzakis, E.; Schaffner, S. F.; Yu, F.; Bonnen, P. E.; de Bakker, P. I.; Deloukas, P.; Gabriel, S. B.; Gwilliam, R.; Hunt, S.; Inouye, M.; Jia, X.; Palotie, A.; Parkin, M.; Whittaker, P.; Chang, K.; Hawes, A.; Lewis, L. R.; Ren, Y.; Wheeler, D.; Muzny, D. M.; Barnes, C.; Darvishi, K.; Hurles, M.; Korn, J. M.; Kristiansson, K.; Lee, C.; McCarrol, S. A.; Nemesh, J.; Keinan, A.; Montgomery, S. B.; Pollack, S.; Price, A. L.; Soranzo, N.; Gonzaga-Jauregui, C.; Anttila, V.; Brodeur, W.; Daly, M. J.; Leslie, S.; McVean, G.; Moutsianas, L.; Nguyen, H.; Zhang, Q.; Ghori, M. J.; McGinnis, R.; McLaren, W.; Takeuchi, F.; Grossman, S. R.; Shlyakhter, I.; Hostetter, E. B.; Sabeti, P. C.; Adebamowo, C. A.; Foster, M. W.; Gordon, D. R.; Licinio, J.; Manca, M. C.; Marshall, P. A.; Matsuda, I.; Ngare, D.; Wang, V. O.; Reddy, D.; Rotimi, C. N.; Royal, C. D.; Sharp, R. R.; Zeng, C.; Brooks, L. D.; McEwen, J. E., Integrating common and rare genetic variation in diverse human populations. *Nature,* **2010**, *467* (7311), 52-8.

[32] Wang, X.; Elston, R. C.; Zhu, X., The meaning of interaction. *Hum. Hered.,* **2010**, *70* (4), 269-77.

[33] Pomerleau, O. F.; Collins, A. C.; Shiffman, S.; Pomerleau, C. S., Why some people smoke and others do not: new perspectives. *J. Consult Clin. Psychol.,* **1993**, *61* (5), 723-31.

[34] Hudmon, K. S.; Marks, J. L.; Pomerleau, C. S.; Bolt, D. M.; Brigham, J.; Swan, G. E., A multidimensional model for characterizing tobacco dependence. *Nicotine Tob. Res.,* **2003**, *5* (5), 655-64.

[35] Moolchan, E. T.; Radzius, A.; Epstein, D. H.; Uhl, G.; Gorelick, D. A.; Cadet, J. L.; Henningfield, J. E., The Fagerstrom Test for Nicotine Dependence and the Diagnostic Interview Schedule: do they diagnose the same smokers? *Addict Behav.,* **2002**, *27* (1), 101-13.

[36] Lombardo, T. W.; Hughes, J. R.; Fross, J. D. Failure to support the validity of the Fagerstrom Tolerance Questionnaire as a measure of physiological tolerance to nicotine. *Addict Behav.,* **1988**, *13* (1), 87-90.

[37] Colby, S. M.; Tiffany, S. T.; Shiffman, S.; Niaura, R. S. Measuring nicotine dependence among youth: a review of available approaches and instruments. *Drug Alcohol Depend.,* **2000**, *59 Suppl 1*, S23-39.

[38] Kandel, D.; Schaffran, C.; Griesler, P.; Samuolis, J.; Davies, M.; Galanti, R. On the measurement of nicotine dependence in adolescence: comparisons of the mFTQ and a DSM-IV-based scale. *J. Pediatr. Psychol.,* **2005**, *30* (4), 319-32.

[39] Lu, Q.; Wei, C.; Ye, C.; Li, M.; Elston, R. C. A Likelihood Ratio-Based Mann-Whitney Approach Finds Novel Replicable Joint Gene Action for Type 2 Diabetes. *Genet. Epidemiol.,* **2012**, *36*(6),583-93

[40] Bettler, B.; Kaupmann, K.; Mosbacher, J.; Gassmann, M. Molecular structure and physiological functions of GABA(B) receptors. *Physiol. Rev.,* **2004**, *84* (3), 835-67.

[41] Beuten, J.; Ma, J. Z.; Payne, T. J.; Dupont, R. T.; Crews, K. M.; Somes, G.; Williams, N. J.; Elston, R. C.; Li, M. D. Single- and multilocus allelic variants within the GABA(B) receptor subunit 2 (GABAB2) gene are significantly associated with nicotine dependence. *Am. J. Hum. Genet.,* **2005**, *76* (5), 859-64.

[42] Li, M. D.; Payne, T. J.; Ma, J. Z.; Lou, X. Y.; Zhang, D.; Dupont,

R. T.; Crews, K. M.; Somes, G.; Williams, N. J.; Elston, R. C. A genomewide search finds major susceptibility loci for nicotine dependence on chromosome 10 in African Americans. *Am. J. Hum. Genet.,* **2006**, *79* (4), 745-51.

[43] Li, M. D.; Ma, J. Z.; Payne, T. J.; Lou, X. Y.; Zhang, D.; Dupont, R. T.; Elston, R. C. Genome-wide linkage scan for nicotine dependence in European Americans and its converging results with African Americans in the Mid-South Tobacco Family sample. *Mol. Psychiatry,* **2008**, *13* (4), 407-16.

[44] Agrawal, A.; Pergadia, M. L.; Saccone, S. F.; Hinrichs, A. L.; Lessov-Schlaggar, C. N.; Saccone, N. L.; Neuman, R. J.; Breslau, N.; Johnson, E.; Hatsukami, D.; Montgomery, G. W.; Heath, A. C.; Martin, N. G.; Goate, A. M.; Rice, J. P.; Bierut, L. J.; Madden, P. A. Gamma-aminobutyric acid receptor genes and nicotine dependence: evidence for association from a case-control study. *Addiction,* **2008**, *103* (6), 1027-38.

[45] Li, M. D.; Mangold, J. E.; Seneviratne, C.; Chen, G. B.; Ma, J. Z.; Lou, X. Y.; Payne, T. J. Association and interaction analyses of GABBR1 and GABBR2 with nicotine dependence in European- and African-American populations. *PLoS One,* **2009**, *4* (9), e7055.

[46] Stopkova, P.; Saito, T.; Fann, C. S.; Papolos, D. F.; Vevera, J.; Paclt, I.; Zukov, I.; Stryjer, R.; Strous, R. D.; Lachman, H. M. Polymorphism screening of PIP5K2A: a candidate gene for chromosome 10p-linked psychiatric disorders. *Am. J. Med. Genet. B Neuropsychiatr. Genet.,* **2003**, *123B* (1), 50-8.

[47] Schwab, S. G.; Knapp, M.; Sklar, P.; Eckstein, G. N.; Sewekow, C.; Borrmann-Hassenbach, M.; Albus, M.; Becker, T.; Hallmayer, J. F.; Lerer, B.; Maier, W.; Wildenauer, D. B. Evidence for association of DNA sequence variants in the phosphatidylinositol-4-phosphate 5-kinase IIalpha gene (PIP5K2A) with schizophrenia. *Mol. Psychiatry,* **2006**, *11* (9), 837-46.

[48] Thiselton, D. L.; Maher, B. S.; Webb, B. T.; Bigdeli, T. B.; O'Neill, F. A.; Walsh, D.; Kendler, K. S.; Riley, B. P. Association analysis of the PIP4K2A gene on chromosome 10p12 and schizophrenia in the Irish study of high density schizophrenia families (ISHDSF) and the Irish case-control study of schizophrenia (ICCSS). *Am. J. Med. Genet. B Neuropsychiatr. Genet.,* **2010**, *153B* (1), 323-31.

[49] Rethelyi, J. M.; Bakker, S. C.; Polgar, P.; Czobor, P.; Strengman, E.; Pasztor, P. I.; Kahn, R. S.; Bitter, I. Association study of NRG1, DTNBP1, RGS4, G72/G30, and PIP5K2A with schizophrenia and symptom severity in a Hungarian sample. *Am. J. Med. Genet. B Neuropsychiatr. Genet.,* **2010**, *153B* (3), 792-801.

[50] Saccone, S. F.; Hinrichs, A. L.; Saccone, N. L.; Chase, G. A.; Konvicka, K.; Madden, P. A.; Breslau, N.; Johnson, E. O.; Hatsukami, D.; Pomerleau, O.; Swan, G. E.; Goate, A. M.; Rutter, J.; Bertelsen, S.; Fox, L.; Fugman, D.; Martin, N. G.; Montgomery, G. W.; Wang, J. C.; Ballinger, D. G.; Rice, J. P.; Bierut, L. J. Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. *Hum. Mol. Genet.,* **2007**, *16* (1), 36-49.