

# Family based association study with complex survey data

DEWEI SHE, HONG ZHANG, YAN LI,  
BARRY I. GRAUBARD, AND ZHAOHAI LI\*

---

Genetic data collected from the Third National Health and Nutrition Examination Survey (NHANES III) provides an opportunity to investigate associations between genetic variations and health-related phenotypes for the US population. Complex sample designs involving stratified multistage cluster sampling and sample weighting are used to sample families in household surveys such as the NHANES III. We modified conditional likelihood score and trend tests used to test the null hypothesis of no association between a candidate gene and a phenotype in simple random samples of nuclear families so that these tests are applicable to data from complex sample designs. The finite sample properties of our modified test procedures are evaluated via Monte Carlo simulation studies. We recommend using an F-version of the trend test instead of a score test because the F-test shows greater power. Our test statistics are applied to NHANES III data to test for associations between the locus *ADRB2* (rs1042713) and obesity, *VDR* (rs2239185) and high blood lead level, and *TGFB1* (rs1982073) and asthma.

KEYWORDS AND PHRASES: Complex sampling, Conditional likelihood score test, Nuclear family, Survey data, Trend test.

---

## 1. INTRODUCTION

Genetic association studies can be classified into two categories, i.e., population-based studies and family-based studies. Population-based studies use genetic information from unrelated individuals while family-based studies employ information from nuclear families or pedigrees. Each of these study designs has advantages and disadvantages. Generally speaking, population-based study designs provide greater power and are easier to implement when compared with family-based studies; however, the population-based study design can be subject to confounding such as population stratification. When population substructure exists, i.e., the allele frequencies vary across sub-populations, and in conjunction with varying disease prevalence rates, a population-based study design may lead to a spurious disease-gene association (Li, 1969) (Lander and Schork, 1994) (Ewens and

Spielman, 1995) (Cardon and Palmer, 2003) (Hattersley and McCarthy, 2005) (Li *et al.*, 2005) (Wang *et al.*, 2005) (Evangelou *et al.*, 2006). Family-based study designs are more robust to population stratification, but they require more effort in terms of designing the study and collecting the samples.

Family-based designs use controls from the same family as cases, as opposed to unrelated controls in a population-based study design. Even though some family-based association tests do not require parental genetic information, many test procedures are based on nuclear families with genetic marker information from parents and their offspring (Li *et al.*, 2005) (Abecasis *et al.*, 2000a) (Rakovski *et al.*, 2007) (Thornton and McPeck, 2010). A prime example of a family-based association test is the transmission disequilibrium test (TDT) (Spielman *et al.*, 1993) (Rabinowitz, 1997) (Risch and Teng, 1998) (Abecasis *et al.*, 2000b). The original TDT test uses data on transmission of marker alleles from heterozygous parents to affected offspring. Li *et al.* (2005) proposed conditional likelihood score tests for association studies in nuclear families with parental genotypes (The conditional likelihood function is the conditional probability of offspring genotypes given the disease status of offspring and parental mating type. The conditional likelihood score test was derived by taking the derivative of the log conditional likelihood function). Under the additive model, the conditional score test is equivalent to the TDT while this is not the case for dominant or recessive models.

The Third National Health and Nutrition Examination Survey (NHANES III) collected blood lymphocytes from 7,159 participants age 12 years and older in its second phase (from October, 1991 to October, 1994) in anticipation of advances in genetic research (NCHS, 2008). Since NHANES III is a nationally representative cross-sectional household survey that collects a multitude of health behavior and medical examination measurements, linkage of the NHANES III health-related phenotype data to its genetic information provides an opportunity to investigate genetic associations with a wide variety of health factors. However, household surveys such as NHANES III usually have stratified multistage cluster sample designs, which can add complexities to data analyses. For the NHANES III at the first stage of the sampling, the target population is partitioned into primary sampling units (PSU), e.g., counties. The PSUs are

---

\*Corresponding author.

then grouped into strata. Within each stratum, PSUs are sampled with probability proportional to the population size. At the second stage of sampling the sampled PSUs are subdivided into smaller geographic units, such as city blocks, and a random sample of these units is selected. Additional stages of sampling may be used in some surveys where smaller units are sampled from the sampled units at the second stage. For the NHANES III, at the third stage of sampling, households (which are often approximately nuclear families) are sampled and finally individuals are sampled from the sampled households. For each individual, the inverse of the product of the selection probabilities across all of the stages of sampling (along with adjustments for non-response and for post-stratification to population totals) is their sample weight. The disease status is ascertained in the sampled offspring as well as among the sampled parents.

Test procedures developed for simple random samples are generally unsuitable for the analysis of data from these complex sample designs such as used in NHANES III. Even though family-based designs with corresponding analytic methods such as TDT are robust to population stratification, these procedures do not address within-family genetic associations with phenotypes that vary across families as a result of possible epigenetic processes or interactions with unknown environmental variables. If the sample weights for the households/families are correlated with the within-family genetic associations, then sample weighted analyses may differ from unweighted analyses. In such cases, sample weighted analyses are preferable as they will provide consistent estimates of genetic association for the target population that will be reproducible across different sample designs applied to the same target population. In addition, as shown in this paper, family-based test statistics that are appropriate for simple random samples can have inflated Type I error when applied to multistage cluster samples. This occurs because of increased variances resulting from correlation of within-family genetic associations among families from the same sampled PSUs (intra-cluster correlation). She *et al.* (2009a) modified several trend tests for genetic association for population-based designs that can appropriately use complex survey data. However, their methods ignore possible family structure information where individuals from the same sampled household are in the sample. The objectives of this paper is to 1) modify test procedures for complex sample designs that can be used to test for an association between a candidate gene and a health related phenotype using a family-based study design rather than a population-based design and 2) use simulations to study the performance of test procedures for family-based studies and compare them to test procedures for population-based studies when the samples are complex multistage stratified cluster samples similar to those from household surveys.

## 2. METHOD

### 2.1 Conditional likelihood score test

Under simple random sampling, let allele  $A$  be the allele of high risk and allele  $a$  be any of the other alleles. Suppose we have a sample of  $n$  nuclear families. For the  $i^{th}$  family, let the number of affected and unaffected sibs with 0, 1, and 2 alleles of  $A$  be  $(j_{i0}, j_{i1}, j_{i2})$  and  $(k_{i0}, k_{i1}, k_{i2})$  respectively and  $j_{i0} + j_{i1} + j_{i2} = r_i$  and  $k_{i0} + k_{i1} + k_{i2} = s_i$ . Denote  $K$  as the disease prevalence,  $\mathbf{f} = (f_0, f_1, f_2)^T$  as the penetrances of genotypes  $(aa, aA, AA)$ ,  $T_{G0} = p(aa|M)$ ,  $T_{G1} = p(aA|M)$ , and  $T_{G2} = p(AA|M)$ , where  $M = (ij)$  is the mating type with  $i$  and  $j$  being the number of  $A$  alleles of the parents. Note that  $K = \sum_{i=0}^2 f_i g_i$  where  $\mathbf{g} = (g_0, g_1, g_2)^T$  are the population genotype frequencies. The conditional likelihood score test statistic for the null hypothesis of no association between the candidate gene and the disease of interest, i.e.,  $H_0 : f_0 = f_1 = f_2$ , given by Li *et al.* (2005) is:  $T_{CLS} = \frac{U}{\sqrt{\widehat{Var}(U)}}$  or  $T_{CLS}^2$ , where  $U = \sum_{i=1}^n U_i$  is the conditional likelihood score, and

$$(1) \quad U_i = \frac{1}{f_0} \sum_{m=0}^2 x_m (j_{im} - r_i T_{Gm}) - \frac{1}{1 - f_0} \sum_{m=0}^2 x_m (k_{im} - s_i T_{Gm}).$$

The score vector  $\mathbf{x} = (x_0, x_1, x_2)^T$  assigned to genotypes  $(aa, aA, AA)$  has the value  $(0, 0.5, 1)^T$  for the additive model,  $(0, 0, 1)^T$  for the recessive model, and  $(0, 1, 1)^T$  for the dominant model.

Under complex sampling, the weighted conditional likelihood score is  $U_w = \sum_{i=1}^n w_i U_i$ , where  $w_i$  is the sample weight associated with the  $i^{th}$  family. (Throughout the Methods Section we will assume that families are sampled and all members of each sampled family are observed. In the Discussion we will address sampling family members.) Let  $z^{(hl)} = \sum_{i=1}^n w_i U_i \delta_i^{(hl)}$  be the weighted sum of  $U_i$ 's for the  $l^{th}$  sampled PSU of the  $h^{th}$  stratum, where  $\delta_i^{(hl)} = 1$  if the  $i^{th}$  family belongs to the  $l^{th}$  sampled PSU of the  $h^{th}$  stratum and 0 otherwise. The variance of  $U_w$  can be estimated by  $\widehat{Var}(U_w) = \sum_{h=1}^H \frac{m_h}{m_h - 1} \sum_{l=1}^{m_h} (z^{(hl)} - \bar{z}^{(h)})^2$ , where  $\bar{z}^{(h)} = \sum_{l=1}^{m_h} \frac{z^{(hl)}}{m_h}$ ,  $H$  is the total number of strata, and  $m_h$  is the number of sampled PSUs within the  $h^{th}$  stratum (Korn and Graubard, 1999). Under  $H_0$ ,  $T_{CLS_w} = \frac{U_w}{\sqrt{\widehat{Var}(U_w)}}$  has an asymptotic normal distribution with mean 0 and variance 1 and  $T_{CLS_w}^2$  has an asymptotic  $\chi_1^2$  distribution (Graubard and Korn, 1993). When  $\nu = \sum_{h=1}^H m_h - H$  is not large, we can use an  $F$ -version of the test statistic  $FT_{CLS_w}^2 = T_{CLS_w}^2$ , to take account of the variability of the denominator of  $T_{CLS_w}^2$  and approximate the distribution of  $FT_{CLS_w}^2$  as a  $F$  distribution with degrees of freedom 1 and  $\nu$  under  $H_0$  (Korn and Graubard, 1990) (Li and Graubard, 2009) (Equivalently a  $t$ -distribution with  $\nu$  degrees of freedom can be used as the reference distribution for  $T_{CLS_w}$  to test  $H_0$ ).

The unweighted or weighted conditional likelihood score involves the population penetrance of genotype  $aa$ ,  $f_0$ , which can either be obtained from another study or be estimated in the current study. Under simple random sampling,  $f_0$  can be estimated as the number of diseased subjects with genotype  $aa$  divided by the total number of subjects with genotype  $aa$ . Under complex sampling with sample weighting,  $f_0$  can be estimated as the sample weighted number of diseased subjects with genotype  $aa$  divided by total sample weighted number of subjects with genotype  $aa$ . When  $f_0$  is estimated, the variance estimator of the conditional score needs to be modified to take this into account. This ratio estimator of  $f_0$  is approximately unbiased under the complex sampling design for the selection of parents and children that we considered because the ascertainment of disease status is determined after the sample is selected. Unlike in the case-control study design, the disease status in the cross-sectional survey design is random instead of fixed.

## 2.2 Trend test

For the  $i^{\text{th}}$  family, let  $X_{ij}(Y_{ij})$  denote the number of  $A$  alleles of the  $j^{\text{th}}$  affected (unaffected) sib and  $X_{if}(X_{im})$  denote the number of  $A$  alleles of father (mother).  $\delta_{if} = 1$  ( $\delta_{im} = 1$ ) if the father (mother) is affected and 0 if unaffected.  $G_{if} = 1$  ( $G_{im} = 1$ ) if the genotype of the father (mother) is known and 0 otherwise. Under a simple random sample of families a trend score can be defined as  $Z_{TT} = \sum_{i=1}^n Q_i$ , where (for notational simplicity, the index  $i$  is omitted when we discuss the trend score and its variance for the  $i^{\text{th}}$  family)

$$Q = \sum_{k=0}^2 x_k \left( \frac{\sum_{j=1}^r I_{\{x_j=k\}} + \delta_f G_f I_{\{x_f=k\}} + \delta_m G_m I_{\{x_m=k\}}}{r + \delta_f G_f + \delta_m G_m} - \frac{\sum_{j=1}^s I_{\{y_j=k\}} + (1 - \delta_f) G_f I_{\{x_f=k\}} + (1 - \delta_m) G_m I_{\{x_m=k\}}}{s + (1 - \delta_f) G_f + (1 - \delta_m) G_m} \right),$$

which is the weighted sum of the frequency difference of each genotype between affected and unaffected family members. Here the score vector  $\mathbf{x} = (x_0, x_1, x_2)^T$  assigned to genotypes ( $aa, aA, AA$ ) is defined the same way as in the previous section. Let  $D(\bar{D})$  be the number of affected (unaffected) family members with available genetic information and  $D_j(\bar{D}_j)$  be the number of affected (unaffected) family members with  $j$  number of allele  $A$ ,  $j = 0, 1, \text{ or } 2$ , then  $Q = \sum_{j=0}^2 x_j \left( \frac{D_j}{D} - \frac{\bar{D}_j}{\bar{D}} \right)$ . Let  $p_j$  be the frequency of genotype with  $j$  number of allele  $A$ ,  $j = 0, 1, \text{ or } 2$ . Under the null hypothesis,

$$E_{H_0}(Q) = \sum_{j=0}^2 x_j \left( \frac{r p_j + \delta_f G_f p_j + \delta_m G_m p_j}{r + \delta_f G_f + \delta_m G_m} - \frac{s p_j + (1 - \delta_f) G_f p_j + (1 - \delta_m) G_m p_j}{s + (1 - \delta_f) G_f + (1 - \delta_m) G_m} \right) = 0$$

The  $E_{H_0}(Z_{TT}) = 0$  and, assuming families are independent of each other, the  $Var_{H_0}(Z_{TT}) = \sum_{i=1}^n Var_{H_0}(Q_i)$ . Under  $H_0$ ,  $T_{TT} = \frac{Z_{TT}}{\sqrt{Var_{H_0}(Z_{TT})}}$  has an asymptotic normal

distribution with mean 0 and variance 1 which follows from the central limit theorem (the  $Q_i$  are independent with finite first and second moments) and  $T_{TT}^2$  has an asymptotic  $\chi_1^2$  distribution. The heuristic justification for using  $T_{TT}^2$  for simple random cross-sectional samples follows from the use of similar test statistics in case-control studies with simple random samples of families (Sasieni, 1997) (Slager and Schaid, 2001) (Freidlin *et al.*, 2002) (Zheng and Gastwirth, 2006). Next we extend  $T_{TT}^2$  to weighted cross-sectional samples with complex sample designs.

In a complex sampling setting, let  $w_i$  be the sample weight associated with the  $i^{\text{th}}$  family.

The total weighted score would be  $Z_{TTw} = \sum_{i=1}^n w_i Q_i$ . When cluster effects exist, the variance of  $Z_{TTw}$  can be estimated using the same procedures described in section 2.1. Under the null hypothesis,  $T_{TTw} = \frac{Z_{TTw}}{\sqrt{Var(Z_{TTw})}}$  has an asymptotic normal distribution with mean 0 and variance 1 and  $T_{TTw}^2$  has an asymptotic  $\chi_1^2$  distribution (Graubard and Korn, 1993). The variance estimate  $\widehat{Var}(Z_{TTw})$  can be derived using the similar Taylor linearization method as discussed in the previous section for  $\widehat{Var}(U_w)$ . When  $v = \sum_{h=1}^H m_h - H$  is not large, we can use an  $F$ -version of the test statistic  $FT_{TTw}^2 = T_{TTw}^2$ , which has an asymptotic  $F$  distribution with degrees of freedom 1 and  $\nu$  under  $H_0$  (Korn and Graubard, 1990) (Li and Graubard, 2009) (She *et al.*, 2009b) (Equivalently a  $t$ -distribution with  $\nu$  degrees of freedom can be used as the reference distribution for  $T_{TTw}$  to test  $H_0$ ).

The conditional likelihood score test requires that each nuclear family has at least one child and has available genetic information for parents and children. On the other hand, the trend test requires that each nuclear family has at least one unaffected member and one affected member with available genetic information. To form the trend test, we restrict to those families having at least one unaffected member and one affected member with genetic information available.

## 2.3 Simulation studies

The type I error rate and power of the modified test procedures are evaluated via Monte Carlo simulation with a two-sided significance level of 0.05. The following parameters were controlled in the same manner used by She *et al.* (2009a): 1)  $H$ , the number of strata; 2)  $L$ , the number of sampled PSUs per stratum; 3)  $\mathbf{x}$ , the score vector assigned to genotypes ( $aa, aA, AA$ ); 4)  $K$ , the population disease prevalence; 5)  $p_A$ , the allele  $A$  frequency; and 6)  $\gamma = (1, \frac{f_1}{f_0}, \frac{f_2}{f_0})^T = (1, \gamma_1, \gamma_2)^T$ , the relative risk. In addition to these parameters, we also controlled 1)  $F$ , the inbreeding coefficient for population substructure or cryptic relatedness, in the same way as by She *et al.* (2009b); and 2) the number of sampled nuclear families per PSU.

We set  $H = 23$  and  $L = 2$  with 60 families per PSU, which is similar to the sample design of the second phase of

NHANES III. The type I error rate and power are evaluated with  $p_A = 0.1, 0.3, \text{ or } 0.5$ . Two different sets of sample weights are considered, i.e.,  $w_j$ 's are equal to one (equivalent to using no weights) or each family is randomly assigned the values  $\{1, 2, 3, 4, 5\}$ , in our simulations. Each family is randomly set to have 0, 1, or 2 sibs. Three score vectors  $\mathbf{x} = (0, 0, 1)^T, (0, 0.5, 1)^T, \text{ and } (0, 1, 1)^T$  are used to generate three different test statistics corresponding to three different genetic models. The population disease prevalence  $K$  is set to be 0.05 or 0.1. The relative risk  $\gamma$  is set to be  $(1, 1, 1)^T$  to evaluate the type I error rate and is set to be  $(1, 1, 1.5)^T, (1, 1.25, 1.5)^T, \text{ and } (1, 1.5, 1.5)^T$  for a recessive model, additive model, and dominant model respectively to evaluate power under alternative hypotheses. The inbreeding coefficient  $F$  was set to be 0 or 0.02.

Genotype data for parents are generated using a two-step procedure assuming observations within PSU are correlated and PSUs are independent of each other. In step 1, the allele  $A$  frequency  $p$  in each PSU was generated from the Beta distribution  $Beta((1 - F)p_A/F; (1 - F)(1 - p_A)/F)$  for inbreeding coefficient  $F > 0$ , and when  $F = 0$  then  $p = p_A$ . In step 2, for each parent, two alleles were drawn at random from the binomial distribution  $Bin(2; p)$  (Zheng *et al.*, 2005). Given parental genotypes, the genotype of a child was randomly generated according to Mendelian law of inheritance. This procedure results in intraclass correlation of the alleles within PSUs, i.e., where the allele frequencies vary across PSUs. We followed the method described in Hunsberger *et al.* (2008) and She *et al.* (2009a), which is based on the method of Oman and Zucker (2000), to generate cases and controls that are correlated within each cluster (PSU) with intra-cluster correlation coefficient of 0.09. To be specific, first we generate  $\varepsilon_{i0}$  for all observations in cluster  $i$  and generate  $\varepsilon_{ij}$  for each observation  $j$  in cluster  $i$ , where  $\varepsilon_{i0}$  and  $\varepsilon_{ij}$  are *iid* from a standard normal distribution. Then we generate  $u_{ij}$  according to a Bernoulli distribution with probability of the square root of intra-cluster correlation coefficient, i.e.  $\sqrt{0.09} = 0.3$ . Let  $\pi_{ij} = f_0 G_j^0 + f_1 G_j^1 + f_2 G_j^2$  be the probability of being a case where  $G_j^i = 1$  if the  $j$ th individual has  $i$  copies of  $A$  alleles at the locus of interest. Observation  $j$  in cluster  $i$  will be defined as a case if  $u_{ij}\varepsilon_{i0} + (1 - u_{ij})\varepsilon_{ij} \leq \phi^{-1}(\hat{\pi}_{ij})$ , otherwise it will be a control.

### 3. RESULT

#### 3.1 Simulation studies

The type I error rate (level) and power of the test procedures are estimated by the proportion of rejection of the null hypothesis at the nominal level of 0.05 in 1,000 independent replications. The binomial standard error for the simulated levels when the true level is  $\alpha = 0.05$  is  $\sqrt{0.05(1 - 0.05)/1000} = 0.00689$ , levels that equal to or exceed  $0.05 + 1.96 \times 0.00689 = 0.064$  are considered to be inflated, otherwise the level is considered to be maintained.

Parameter settings where the nominal level is maintained in Tables 1–5 are bolded. In our simulation, we calculate the test statistics using the true  $f_0$  and  $T_{Gm}$  used for the data generation.

Table 1 displays the type I error rates of the test statistics assuming there is no population substructure and no intraclass correlation in disease status. F-version test statistics maintain the nominal level except for a couple of occasions. Type I error rates are similar when there is population substructure and disease status is correlated within each cluster (table 2). In general,  $FT_{CLS_w}^2$  test and  $FT_{TT_w}^2$  test are superior to their corresponding  $T_{CLS_w}^2$  and  $T_{TT_w}^2$  in terms of maintaining the level of the test (data not shown).

Table 3 presents the power achieved by the test statistics with sample weights when there is no population substructure and no intraclass correlation in disease status. Table 4 presents the power when there is population substructure and disease status is correlated within each cluster where the values of power are bolded for parameter settings in which the nominal level is maintained (see table 2).

Tests are more powerful when the model is correctly specified than when it is incorrectly specified. When the data are generated using recessive or dominant models and incorrect genetic models are used for testing the association, tests assuming an additive model, i.e.,  $FT_{CLS_w}^2(0.5)$  and  $FT_{TT_w}^2(0.5)$ , are more powerful in general than tests that assume incorrect genetic models. All tests tend to have less power when there is population substructure and/or intraclass correlation in disease status compared to when there is no population structure and no intraclass correlation in disease status. The above discussion holds even if the analyses were restricted to families with at least one child.

When analyses were restricted to those families with at least one unaffected member and one affected member,  $FT_{TT_w}^2(0.5)$  has greater power than  $FT_{CLS_w}^2(0.5)$  (in table 5 the values of power are bolded for parameter settings where the nominal level is maintained).

Although not presented, we conducted simulations of the power of the test statistics with no sample weights compared to sample weighted tests when the sample weights are randomly assigned i.e., noninformative weights. In general we found the unweighted tests had greater power, which would be expected since weighting tends to increase variances (Korn and Graubard, 1999).

#### 3.2 Real data analysis

To illustrate our methods, we apply the test statistics to the NHANES III genetic data for the same three candidate loci analyzed by She *et al.* (2009a), i.e., *ADRB2* (rs1042713), *TGFB1* (rs1982073), and *VDR* (rs2239185). In this example we will assume that there is no sub-sampling of family members. The NHANES III does not provide a sample weight for each family. Instead, it provides a sample weight for each participant. We use the approach of She *et al.* (2009b) to calculate a sample weight for each

Table 1. Type I error rates given by test statistics: no population substructure and no intraclass correlation in disease status

K			$w \equiv 1$			$w \in \{1, 2, 3, 4, 5\}$		
			$p_A \equiv 0.1$	0.3	0.5	0.1	0.3	0.5
0.05	Score Test (1 d.f.)	$FT_{CLS_w}^2(0)$	0.092	0.065	<b>0.056</b>	0.070	<b>0.057</b>	<b>0.050</b>
		$FT_{CLS_w}^2(0.5)$	<b>0.058</b>	<b>0.050</b>	<b>0.046</b>	<b>0.047</b>	<b>0.059</b>	<b>0.054</b>
		$FT_{CLS_w}^2(1)$	<b>0.048</b>	<b>0.051</b>	<b>0.051</b>	<b>0.046</b>	<b>0.048</b>	<b>0.050</b>
Trend Test (1 d.f.)	$FT_{TT_w}^2(0)$	<b>0.045</b>	<b>0.041</b>	<b>0.047</b>	<b>0.049</b>	<b>0.049</b>	<b>0.052</b>	
	$FT_{TT_w}^2(0.5)$	<b>0.044</b>	<b>0.049</b>	<b>0.041</b>	<b>0.040</b>	<b>0.054</b>	<b>0.040</b>	
	$FT_{TT_w}^2(1)$	<b>0.041</b>	<b>0.051</b>	<b>0.048</b>	<b>0.046</b>	<b>0.040</b>	<b>0.041</b>	
0.1	Score Test (1 d.f.)	$FT_{CLS_w}^2(0)$	0.068	<b>0.059</b>	<b>0.051</b>	0.075	<b>0.048</b>	<b>0.038</b>
		$FT_{CLS_w}^2(0.5)$	<b>0.060</b>	<b>0.044</b>	<b>0.057</b>	<b>0.042</b>	<b>0.051</b>	<b>0.048</b>
		$FT_{CLS_w}^2(1)$	<b>0.057</b>	<b>0.033</b>	<b>0.048</b>	<b>0.042</b>	<b>0.052</b>	<b>0.060</b>
Trend Test (1 d.f.)	$FT_{TT_w}^2(0)$	<b>0.046</b>	<b>0.049</b>	<b>0.049</b>	0.075	<b>0.048</b>	<b>0.038</b>	
	$FT_{TT_w}^2(0.5)$	<b>0.048</b>	<b>0.044</b>	<b>0.058</b>	<b>0.042</b>	<b>0.051</b>	<b>0.048</b>	
	$FT_{TT_w}^2(1)$	<b>0.050</b>	<b>0.042</b>	<b>0.054</b>	<b>0.042</b>	<b>0.052</b>	<b>0.060</b>	

1.  $FT_{CLS_w}^2(x_1)$  and  $FT_{TT_w}^2(x_1)$ :  $x_1 = 0$  for recessive model;  $x_1 = 0.5$  for additive model;  $x_1 = 1$  for dominant model.

2.  $K$ : the disease prevalence.

3. The null hypothesis is that all penetrances are equal.

4. Simulated data have 23 strata with 2 PSUs per stratum and 60 families per cluster.

5. Relative risk  $(1, 1, 1)^T$  is used.

6.  $w$ : the sample weight for each family.

7. Parameter settings where the nominal level is maintained are bolded.

Table 2. Type I error rates given by test statistics: population substructure exists and disease status is correlated within each cluster

K			$w \equiv 1$			$w \in \{1, 2, 3, 4, 5\}$		
			$p_A \equiv 0.1$	0.3	0.5	0.1	0.3	0.5
0.05	Score Test (1 d.f.)	$FT_{CLS_w}^2(0)$	<b>0.056</b>	<b>0.046</b>	<b>0.043</b>	<b>0.050</b>	<b>0.047</b>	<b>0.048</b>
		$FT_{CLS_w}^2(0.5)$	<b>0.045</b>	<b>0.044</b>	<b>0.036</b>	<b>0.040</b>	<b>0.054</b>	<b>0.047</b>
		$FT_{CLS_w}^2(1)$	<b>0.052</b>	<b>0.053</b>	<b>0.048</b>	<b>0.048</b>	<b>0.057</b>	<b>0.060</b>
Trend Test (1 d.f.)	$FT_{TT_w}^2(0)$	<b>0.040</b>	<b>0.052</b>	<b>0.049</b>	<b>0.047</b>	<b>0.047</b>	<b>0.044</b>	
	$FT_{TT_w}^2(0.5)$	<b>0.051</b>	<b>0.045</b>	<b>0.052</b>	<b>0.041</b>	<b>0.048</b>	<b>0.059</b>	
	$FT_{TT_w}^2(1)$	<b>0.054</b>	<b>0.045</b>	<b>0.050</b>	<b>0.036</b>	<b>0.060</b>	<b>0.051</b>	
0.1	Score Test (1 d.f.)	$FT_{CLS_w}^2(0)$	<b>0.056</b>	<b>0.045</b>	<b>0.052</b>	0.069	<b>0.056</b>	<b>0.047</b>
		$FT_{CLS_w}^2(0.5)$	<b>0.047</b>	<b>0.046</b>	<b>0.054</b>	<b>0.058</b>	<b>0.056</b>	<b>0.040</b>
		$FT_{CLS_w}^2(1)$	<b>0.046</b>	<b>0.048</b>	<b>0.061</b>	<b>0.051</b>	<b>0.063</b>	<b>0.051</b>
Trend Test (1 d.f.)	$FT_{TT_w}^2(0)$	<b>0.036</b>	<b>0.046</b>	<b>0.045</b>	<b>0.044</b>	<b>0.053</b>	<b>0.049</b>	
	$FT_{TT_w}^2(0.5)$	<b>0.052</b>	<b>0.048</b>	<b>0.038</b>	<b>0.055</b>	<b>0.053</b>	<b>0.052</b>	
	$FT_{TT_w}^2(1)$	<b>0.055</b>	<b>0.051</b>	<b>0.054</b>	<b>0.050</b>	<b>0.049</b>	<b>0.059</b>	

1.  $FT_{CLS_w}^2(x_1)$  and  $FT_{TT_w}^2(x_1)$ :  $x_1 = 0$  for recessive model;  $x_1 = 0.5$  for additive model;  $x_1 = 1$  for dominant model.

2.  $K$ : the disease prevalence.

3. The null hypothesis is that all penetrances are equal.

4. Simulated data have 23 strata with 2 PSUs per stratum and 60 families per cluster.

5. Relative risk  $(1, 1, 1)^T$  is used.

6. The inbreeding coefficient  $F$  is 0.02 and the intraclass correlation in disease status is 0.09.

7.  $w$ : the sample weight for each family.

8. Parameter settings where the nominal level is maintained are bolded.

family, i.e., the sample weight for a family is computed as the sum of the individual sample weights from the family members divided by family sample size. For example, for a family of three family members sampled, the sample weight for that family would be the sum of sample weights of the three family members divided by three. The characteristics

of NHANES III data are summarized in Table 6. External  $f_0$ 's (23% for obesity, 7.2% for asthma, and 2.2% for  $BLL \geq 10 \mu\text{g/dL}$ ), which are consistent with the  $f_0$ 's estimated from NHANES III (She *et al.*, 2009b), and  $T_{Gm}$  based on parental genotype are used for conditional likelihood score test.

Table 3. Power given by test statistics: no population substructure, no intraclass correlation in disease status, and  $w \in \{1, 2, 3, 4, 5\}$

$K$		Genetic model used to generate the data									
		Recessive model			Additive model			Dominant model			
		$p_A \equiv 0.1$	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	
0.05	Score Test (1 d.f.)	$FT_{CLS_w}^2(0)$	0.044	<b>0.149</b>	<b>0.273</b>	0.059	<b>0.066</b>	<b>0.091</b>	0.064	<b>0.047</b>	<b>0.044</b>
		$FT_{CLS_w}^2(0.5)$	<b>0.052</b>	<b>0.092</b>	<b>0.197</b>	<b>0.102</b>	<b>0.163</b>	<b>0.180</b>	<b>0.213</b>	<b>0.251</b>	<b>0.156</b>
		$FT_{CLS_w}^2(1)$	<b>0.047</b>	<b>0.051</b>	<b>0.050</b>	<b>0.100</b>	<b>0.165</b>	<b>0.145</b>	<b>0.228</b>	<b>0.327</b>	<b>0.276</b>
Trend Test (1 d.f.)	$FT_{TT_w}^2(0)$	<b>0.073</b>	<b>0.430</b>	<b>0.740</b>	<b>0.058</b>	<b>0.164</b>	<b>0.283</b>	<b>0.046</b>	<b>0.058</b>	<b>0.058</b>	
	$FT_{TT_w}^2(0.5)$	<b>0.050</b>	<b>0.216</b>	<b>0.535</b>	<b>0.236</b>	<b>0.431</b>	<b>0.477</b>	<b>0.571</b>	<b>0.635</b>	<b>0.402</b>	
	$FT_{TT_w}^2(1)$	<b>0.045</b>	<b>0.070</b>	<b>0.073</b>	<b>0.219</b>	<b>0.354</b>	<b>0.349</b>	<b>0.631</b>	<b>0.805</b>	<b>0.669</b>	
0.1	Score Test (1 d.f.)	$FT_{CLS_w}^2(0)$	0.060	<b>0.293</b>	<b>0.528</b>	0.055	<b>0.117</b>	<b>0.162</b>	0.065	<b>0.051</b>	<b>0.049</b>
		$FT_{CLS_w}^2(0.5)$	<b>0.061</b>	<b>0.162</b>	<b>0.367</b>	<b>0.174</b>	<b>0.312</b>	<b>0.319</b>	<b>0.410</b>	<b>0.485</b>	<b>0.293</b>
		$FT_{CLS_w}^2(1)$	<b>0.040</b>	<b>0.059</b>	<b>0.072</b>	<b>0.171</b>	<b>0.280</b>	<b>0.235</b>	<b>0.444</b>	<b>0.645</b>	<b>0.498</b>
Trend Test (1 d.f.)	$FT_{TT_w}^2(0)$	0.151	<b>0.760</b>	<b>0.972</b>	0.086	<b>0.346</b>	<b>0.552</b>	0.060	<b>0.084</b>	<b>0.081</b>	
	$FT_{TT_w}^2(0.5)$	<b>0.070</b>	<b>0.407</b>	<b>0.849</b>	<b>0.454</b>	<b>0.751</b>	<b>0.785</b>	<b>0.881</b>	<b>0.934</b>	<b>0.727</b>	
	$FT_{TT_w}^2(1)$	<b>0.055</b>	<b>0.079</b>	<b>0.116</b>	<b>0.423</b>	<b>0.648</b>	<b>0.594</b>	<b>0.913</b>	<b>0.982</b>	<b>0.924</b>	

- $FT_{CLS_w}^2(x_1)$  and  $FT_{TT_w}^2(x_1)$ :  $x_1 = 0$  for recessive model;  $x_1 = 0.5$  for additive model;  $x_1 = 1$  for dominant model.
- $K$ : the disease prevalence.
- The null hypothesis is that all penetrances are equal.
- The relative risk is set to be  $(1, 1, 1.5)^T$ ,  $(1, 1.25, 1.5)^T$ , and  $(1, 1.5, 1.5)^T$  for recessive model, additive model, and dominant model respectively.
- Simulated data have 23 strata with 2 PSUs per stratum and 60 families per cluster.
- $w$ : the sample weight for each family.
- Parameter settings where the nominal level is maintained are bolded.

Table 4. Power given by test statistics: population substructure exists, disease status is correlated within each cluster, and  $w \in \{1, 2, 3, 4, 5\}$

$K$		Genetic model used to generate the data									
		Recessive model			Additive model			Dominant model			
		$p_A \equiv 0.1$	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	
0.05	Score Test (1 d.f.)	$FT_{CLS_w}^2(0)$	<b>0.035</b>	<b>0.129</b>	<b>0.233</b>	<b>0.047</b>	<b>0.070</b>	<b>0.090</b>	<b>0.052</b>	<b>0.045</b>	<b>0.054</b>
		$FT_{CLS_w}^2(0.5)$	<b>0.048</b>	<b>0.086</b>	<b>0.170</b>	<b>0.088</b>	<b>0.147</b>	<b>0.151</b>	<b>0.177</b>	<b>0.191</b>	<b>0.126</b>
		$FT_{CLS_w}^2(1)$	<b>0.046</b>	<b>0.046</b>	<b>0.072</b>	<b>0.084</b>	<b>0.120</b>	<b>0.138</b>	<b>0.191</b>	<b>0.263</b>	<b>0.220</b>
Trend Test (1 d.f.)	$FT_{TT_w}^2(0)$	<b>0.079</b>	<b>0.358</b>	<b>0.571</b>	<b>0.061</b>	<b>0.157</b>	<b>0.207</b>	<b>0.048</b>	<b>0.058</b>	<b>0.062</b>	
	$FT_{TT_w}^2(0.5)$	<b>0.064</b>	<b>0.200</b>	<b>0.443</b>	<b>0.204</b>	<b>0.348</b>	<b>0.372</b>	<b>0.456</b>	<b>0.493</b>	<b>0.341</b>	
	$FT_{TT_w}^2(1)$	<b>0.043</b>	<b>0.074</b>	<b>0.095</b>	<b>0.188</b>	<b>0.276</b>	<b>0.271</b>	<b>0.478</b>	<b>0.606</b>	<b>0.494</b>	
0.1	Score Test (1 d.f.)	$FT_{CLS_w}^2(0)$	0.049	<b>0.265</b>	<b>0.469</b>	0.038	<b>0.105</b>	<b>0.150</b>	0.047	<b>0.054</b>	<b>0.052</b>
		$FT_{CLS_w}^2(0.5)$	<b>0.072</b>	<b>0.148</b>	<b>0.318</b>	<b>0.173</b>	<b>0.267</b>	<b>0.278</b>	<b>0.361</b>	<b>0.392</b>	<b>0.238</b>
		$FT_{CLS_w}^2(1)$	<b>0.058</b>	<b>0.065</b>	<b>0.061</b>	<b>0.162</b>	<b>0.215</b>	<b>0.209</b>	<b>0.401</b>	<b>0.503</b>	<b>0.414</b>
Trend Test (1 d.f.)	$FT_{TT_w}^2(0)$	<b>0.147</b>	<b>0.676</b>	<b>0.898</b>	<b>0.088</b>	<b>0.289</b>	<b>0.437</b>	<b>0.052</b>	<b>0.084</b>	<b>0.074</b>	
	$FT_{TT_w}^2(0.5)$	<b>0.074</b>	<b>0.394</b>	<b>0.751</b>	<b>0.391</b>	<b>0.642</b>	<b>0.683</b>	<b>0.774</b>	<b>0.828</b>	<b>0.625</b>	
	$FT_{TT_w}^2(1)$	<b>0.045</b>	<b>0.075</b>	<b>0.142</b>	<b>0.371</b>	<b>0.543</b>	<b>0.506</b>	<b>0.810</b>	<b>0.908</b>	<b>0.821</b>	

- $FT_{CLS_w}^2(x_1)$  and  $FT_{TT_w}^2(x_1)$ :  $x_1 = 0$  for recessive model;  $x_1 = 0.5$  for additive model;  $x_1 = 1$  for dominant model.
- $K$ : the disease prevalence.
- The null hypothesis is that all penetrances are equal.
- Simulated data have 23 strata with 2 PSUs per stratum and 60 families per cluster.
- The relative risk is set to be  $(1, 1, 1.5)^T$ ,  $(1, 1.25, 1.5)^T$ , and  $(1, 1.5, 1.5)^T$  for recessive model, additive model, and dominant model respectively.
- The inbreeding coefficient  $F$  is 0.02 and the intraclass correlation in disease status is 0.09.
- $w$ : the sample weight for each family.
- Parameter settings where the nominal level is maintained are bolded.

Table 5. Power given by test statistics: restrict to families with both affected and unaffected, population substructure exists, disease status is correlated within each cluster, and  $w \in \{1, 2, 3, 4, 5\}$

$K$		Genetic model used to generate the data									
		Recessive model			Additive model			Dominant model			
		$p_A \equiv 0.1$	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	
0.05	Score Test (1 d.f.)	$FT_{CLS_w}^2(0)$	<b>0.038</b>	<b>0.129</b>	<b>0.230</b>	<b>0.052</b>	<b>0.063</b>	<b>0.088</b>	<b>0.056</b>	<b>0.055</b>	<b>0.053</b>
		$FT_{CLS_w}^2(0.5)$	<b>0.042</b>	<b>0.089</b>	<b>0.160</b>	<b>0.085</b>	<b>0.144</b>	<b>0.145</b>	<b>0.174</b>	<b>0.197</b>	<b>0.125</b>
		$FT_{CLS_w}^2(1)$	<b>0.040</b>	<b>0.045</b>	<b>0.066</b>	<b>0.085</b>	<b>0.123</b>	<b>0.130</b>	<b>0.180</b>	<b>0.265</b>	<b>0.223</b>
Trend Test (1 d.f.)	$FT_{TT_w}^2(0)$	<b>0.079</b>	<b>0.358</b>	<b>0.571</b>	<b>0.061</b>	<b>0.157</b>	<b>0.207</b>	<b>0.048</b>	<b>0.058</b>	<b>0.062</b>	
	$FT_{TT_w}^2(0.5)$	<b>0.064</b>	<b>0.200</b>	<b>0.443</b>	<b>0.204</b>	<b>0.348</b>	<b>0.372</b>	<b>0.456</b>	<b>0.493</b>	<b>0.341</b>	
	$FT_{TT_w}^2(1)$	<b>0.043</b>	<b>0.074</b>	<b>0.095</b>	<b>0.188</b>	<b>0.276</b>	<b>0.271</b>	<b>0.478</b>	<b>0.606</b>	<b>0.494</b>	
0.1	Score Test (1 d.f.)	$FT_{CLS_w}^2(0)$	<b>0.056</b>	<b>0.251</b>	<b>0.445</b>	<b>0.052</b>	<b>0.102</b>	<b>0.164</b>	<b>0.054</b>	<b>0.047</b>	<b>0.038</b>
		$FT_{CLS_w}^2(0.5)$	<b>0.061</b>	<b>0.144</b>	<b>0.317</b>	<b>0.160</b>	<b>0.263</b>	<b>0.279</b>	<b>0.334</b>	<b>0.383</b>	<b>0.234</b>
		$FT_{CLS_w}^2(1)$	<b>0.047</b>	<b>0.061</b>	<b>0.070</b>	<b>0.149</b>	<b>0.215</b>	<b>0.214</b>	<b>0.374</b>	<b>0.497</b>	<b>0.400</b>
Trend Test (1 d.f.)	$FT_{TT_w}^2(0)$	<b>0.147</b>	<b>0.676</b>	<b>0.898</b>	<b>0.088</b>	<b>0.289</b>	<b>0.437</b>	<b>0.052</b>	<b>0.084</b>	<b>0.074</b>	
	$FT_{TT_w}^2(0.5)$	<b>0.074</b>	<b>0.394</b>	<b>0.751</b>	<b>0.391</b>	<b>0.642</b>	<b>0.683</b>	<b>0.774</b>	<b>0.828</b>	<b>0.625</b>	
	$FT_{TT_w}^2(1)$	<b>0.045</b>	<b>0.075</b>	<b>0.142</b>	<b>0.371</b>	<b>0.543</b>	<b>0.506</b>	<b>0.810</b>	<b>0.908</b>	<b>0.821</b>	

1.  $FT_{CLS_w}^2(x_1)$  and  $FT_{TT_w}^2(x_1)$ :  $x_1 = 0$  for recessive model;  $x_1 = 0.5$  for additive model;  $x_1 = 1$  for dominant model.

2.  $K$ : the disease prevalence.

3. The null hypothesis is that all penetrances are equal.

4. Simulated data have 23 strata with 2 PSUs per stratum and 60 families per cluster.

5. The relative risk is set to be  $(1, 1, 1.5)^T$ ,  $(1, 1.25, 1.5)^T$ , and  $(1, 1.5, 1.5)^T$  for recessive model, additive model, and dominant model respectively.

6. The inbreeding coefficient  $F$  is 0.02 and the intraclass correlation in disease status is 0.09.

7.  $w$ : the sample weight for each family.

8. Parameter settings where the nominal level is maintained are bolded.

Table 6. Characteristics of NHANES III

	<i>ADRB2</i> /Obesity	<i>TGFB1</i> /Asthma	<i>VDR</i> /BLL
Families with two parents & at least one child	150	154	142
Families with affected member	1534	478	188
Families with unaffected member	3656	4302	4368
Families with both affected and unaffected members	699	282	99

Table 7 displays the p-values of the F-version of the conditional likelihood score and trend tests for associations between *ADRB2* and obesity, between *TGFB1* and asthma, and between *VDR* and blood lead level. No significant associations are found for the conditional likelihood score tests between *ADRB2* (rs1042713) and obesity and between *TGFB1* (rs1982073) and asthma. Trend test assuming a recessive model suggests that there is a significant association between *VDR* (rs2239185) and higher blood lead level.

SAS computer code for computing the conditional likelihood score and trend tests is available from the Dewei She upon request.

#### 4. DISCUSSION AND CONCLUSION

In this paper, we modified the conditional likelihood score test and trend test, which are used to test the null hypothesis of no association between a candidate gene and a phenotype in simple random samples of nuclear families, so that these tests can be applied to complex survey

data. The type I error rates and powers of all test statistics are evaluated for finite samples via Monte Carlo simulations with nominal level of 0.05. Type I error rates are well controlled for F-version test statistics (around 5% of the type I errors). Test statistics with the correct genetic model achieve greater power. When the wrong genetic model is assumed, the  $FT_{CLS_w}^2(0.5)$  and  $FT_{TT_w}^2(0.5)$  tests, in general, have greater power than tests based on recessive or dominant models. Compared to the  $FT_{CLS_w}^2(0.5)$  test, the  $FT_{TT_w}^2(0.5)$  test achieves much greater power, which may partially be explained by the fact that the conditional likelihood score test requires that each family has at least one child whereas about one-third of the families in our simulations do not have any children and are not used in the analyses. After restricting the analysis to families with at least one child in our simulation study, the  $FT_{TT_w}^2(0.5)$  test is still more powerful than the  $FT_{CLS_w}^2(0.5)$  test, which might be due to the fact that the  $FT_{CLS_w}^2(0.5)$  test does not use parental disease information and it is conditioned

Table 7. *p*-values for association test for loci *ADRB2*, *TGFB1*, and *VDR*

		<i>ADRB2</i>	<i>TGFB1</i>	<i>VDR</i>
Score Test (1 d.f.)	$FT_{CLS_w}^2(0)$	0.639	0.696	0.606
	$FT_{CLS_w}^2(0.5)$	0.754	0.820	0.778
	$FT_{CLS_w}^2(1)$	0.860	0.987	0.776
Trend Test (1 d.f.)	$FT_{TT_w}^2(0)$	0.092	0.467	0.014
	$FT_{TT_w}^2(0.5)$	0.125	0.401	0.058
	$FT_{TT_w}^2(1)$	0.400	0.539	0.305

on parental genetic information. Therefore, we recommend using the  $FT_{TT_w}^2(0.5)$  to test the association between a disease and a candidate gene in the complex sampling setting with family data available when the genetic model is unknown.

Unlike the tests considered in this paper, the test procedures recommended by She *et al.* (2009a), i.e.,  $FT_{S1w}^2(0.5)$  and  $FT_{T_w}^2(0.5)$ , ignore the family structure. In our simulation study,  $FT_{S1w}^2(0.5)$  and  $FT_{T_w}^2(0.5)$  have greater power than  $FT_{TT_w}^2(0.5)$ . Only when the analyses are restrict to families with at least one unaffected member and one affected member and when there is population substructure and intra-class correlation or the sample weights vary,  $FT_{TT_w}^2(0.5)$  has greater power than the  $FT_{S1w}^2(0.5)$  and  $FT_{T_w}^2(0.5)$  tests (data not shown), in general.

The power of simulated studies depends on 1)  $p_A$ , the allele *A* frequency and 2)  $K$ , the population disease prevalence. When  $p_A$  is small, very few family members have genotype of *AA*. When  $K$  is small, the vast majority of family members are controls. The power for either case is small.

Risch and Teng (1998) proposed a test statistic of the form  $T_{RT} = \frac{\hat{p}_1 - \hat{p}_2}{\hat{\sigma}}$ , in which the numerator is the difference of the estimated allele *A* frequencies between the affected and unaffected groups and the denominator is the estimated standard deviation of the numerator. In the complex sampling setting,  $\hat{p}_1 - \hat{p}_2 = \frac{1}{\sum_{i=1}^n w_i} \sum_{i=1}^n w_i Q_i$ , where  $Q_i$  is the trend score for the  $i^{th}$  family assuming an additive model. It can be shown that this statistic is the same as the trend test statistic assuming an additive model. Other family-based association tests such as sibling-based transmission disequilibrium tests and reconstruction-combined TDT are used for simple random samples. Extending these to complex samples is an area of future research.

When developing our testing procedures, we assumed that all family members are sampled once a family is sampled, which may not always be true. Define  $w_{qi}$  to be the sample weight of sampled individual  $q$  conditional on family  $i$  being sampled, i.e., the reciprocal of the probability of sampling individual  $q$  from family  $i$ . For affected individuals in sampled family  $i$ , let  $\delta_{mqi} = 1$  if the individual is affected and = 0 otherwise, where  $m$  is the number of *A* alleles (0, 1, or 2). Denote  $\hat{j}_{im} = \sum_{q=1}^{n_i} w_{qi} \delta_{mqi}$  and  $\hat{r}_i = \sum_{m=0}^2 \hat{j}_{im}$  and similarly denote  $\hat{k}_{im}$  and  $\hat{s}_i$ , the conditional likelihood score

test can be extended to the setting of differential weighting within each family by replacing  $j_{im}$ ,  $k_{im}$ ,  $r_i$ , and  $s_i$  with  $\hat{j}_{im}$ ,  $\hat{k}_{im}$ ,  $\hat{r}_i$ , and  $\hat{s}_i$  respectively in equation 1. The conditional likelihood score test requires that each family has at least one child and has available genetic information available for parents and children. Differential weighting within each family is for children only. The Cochran-Armitage trend test has been applied to analysis of genetic data under simple random sampling without sample weights (Slager and Schaid, 2001) (Freidlin *et al.*, 2002) (Zheng and Gastwirth, 2006). Properties of trend test for genetic analysis, such as power, sample size, robustness, and estimation of variance have been investigated and reported before (Slager and Schaid, 2001) (Freidlin *et al.*, 2002) (Zheng and Gastwirth, 2006). We adapted trend tests for family based association study with complex survey data. The trend test uses parental genotypes and requires that each nuclear family has both at least one unaffected member and one affected member with available genetic information, which can be extended to the setting of differential weighting within each family in a similar fashion as described above. Even though using parental genotypes in the trend test is not typically done, it makes more complete use of the family data which should result in greater statistical power.

For the purpose of illustration, the test statistics described in this paper are applied to NHANES III data, which is the only survey data with genetic information available that we are aware of, to test for associations between the locus *ADRB2* (rs1042713) and obesity, between *TGFB1* (rs1982073) and asthma, and between *VDR* (rs2239185) and high blood lead level. Unlike She *et al.* (2009a), no significant association is found between *TGFB1* (rs1982073) and asthma. Trend test assuming a recessive model suggests that there is a significant association between *VDR* (rs2239185) and high blood lead level. Since familial relationships were not collected in the NHANES III (the exact relationship among household members is unknown), we inferred the relationships based on age and gender of the surveyed individuals. Families with more than two adult family members or families having two adult family members of the same gender were removed from the analysis. It should be noted that Katki *et al.* (2010) has recently developed a more sophisticated method to infer familial relationships in the NHANES, which could potentially be utilized in family-based association analyses.

## ACKNOWLEDGEMENTS

The authors thank Mr. Christopher Sanders and Dr. Geraldine McQuillan from National Center for Health Statistics, Centers for Disease Control and Prevention for preparing the NHANES III genetic data for us. The work was partially supported by an NIH grant (No. RO1ES016626, Zhaohai Li) and an NSFC grant (No. 11371101, Hong Zhang and Zhaohai Li).



## CONFLICT OF INTEREST

The authors have declared no conflict of interest.

Received 9 May 2013

## REFERENCES

- [1] ABECASIS, G. R., CARDON, L. R., and COOKSON, W. O. (2000a). A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* **66**, 279–292.
- [2] ABECASIS, G. R., COOKSON, W. O., and CARDON, L. R. (2000b). Pedigree tests of transmission disequilibrium. *Eur J Hum Genet* **8**, 545–551.
- [3] CARDON, L. R. and PALMER L. J. (2003). Population stratification and spurious allelic association. *Lancet* **361**, 598–604.
- [4] EVANGELOU, E., TRIKALINOS, T. A., SALANTI, G., and IOANNIDIS, J. P. A. (2006). Family-based versus unrelated case-control designs for genetic associations. *PLoS Genet* **2**, e123.
- [5] EWENS W. and SPIELMAN, R. (1995). The transmission/disequilibrium test: History, subdivision, and admixture. *Am J Hum Genet* **57**, 455–464.
- [6] FREIDLIN, B., ZHENG, G., LI, Z., and GASTWIRTH, J. L. (2002). Trend tests for case-control studies of genetic markers: Power, sample size and robustness. *Hum Hered* **53**, 146–152.
- [7] GRAUBARD, B. I., and KORN, E. L. (1993). Hypothesis testing with complex survey data: The use of classical quadratic test statistics with particular reference to regression problems. *J Am Statist Assoc* **88**, 629–641. [MR1224389](#)
- [8] HATTERSLEY A. T. and MCCARTHY M. (2005). What makes a good genetic association. *Lancet* **366**, 1315–1323.
- [9] HUNSBERGER, S., GRAUBARD, B. I., and KORN, E. L. (2008). Testing logistic regression coefficients with clustered data and few positive outcomes. *Statistics in Medicine* **27**, 1305–1324. [MR2420159](#)
- [10] KATKI, H. A., SANDERS, C. L., GRAUBARD, B. I., and BERGEN A. W. (2010). Using DNA fingerprints to infer familial relationships within NHANES III households. *J Am Statist Assoc* **105**, 552–563. [MR2759931](#)
- [11] KORN, E. L. and GRAUBARD, B. I. (1990). Simultaneous testing of regression coefficients with complex survey data: Use of Bonferroni t statistics. *American Statistician* **44**, 270–276.
- [12] KORN, E. L. and GRAUBARD, B. I. (1999). *Analysis of Health Surveys*. New York, John Wiley & Sons, Inc.
- [13] LANDER, E. S. and SCHORK, N. J. (1994). Genetic dissection of complex traits. *Science* **265**, 2037–2048.
- [14] LI, C. C. (1969). Population subdivision with respect to multiple alleles. *Am J Hum Genet* **33**, 23–29.
- [15] LI, Y. and GRAUBARD, B. I. (2009). Testing Hardy-Weinberg equilibrium and homogeneity of Hardy-Weinberg disequilibrium using complex survey data. *Biometrics* **65**, 1096–1104. [MR2756497](#)
- [16] LI, Z., GASTWIRTH, J. L., and GAIL, M. H. (2005). Power and related statistical properties of conditional likelihood score tests for association studies in nuclear families with parental genotypes. *Ann Hum Genet* **69**, 296–314.
- [17] NCHS. (2008). <http://www.cdc.gov/nchs/about/major/nhanes/nh3datagenetic.htm>.
- [18] OMAN, S. D. and ZUCKER, D. M. (2000). Modeling and generating correlated binary variables. *Biometrika* **88**, 287–290. [MR1841276](#)
- [19] RABINOWITZ, D. (1997). A transmission disequilibrium test for quantitative trait loci. *Hum Hered* **47**, 342–350.
- [20] RAKOVSKI, C., XU, X., LAZARAS, R., and LAIRD N. (2007). A new multimarker test for family-based association studies. *Genet Epidemiol* **31**, 9–17.
- [21] RISCH, N. and TENG, J. (1998). The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human disease I. DNA pooling. *Genome Research* **8**, 1273–1288.
- [22] SASIENI, P. D. (1997). From genotypes to genes: Doubling the sample size. *Biometrics* **53**, 1253–1261. [MR1614374](#)
- [23] SHE, D., LI, Y., ZHANG, H., GRAUBARD, B. I., and LI, Z. (2009a). Trend tests for genetic association using population based cross-sectional complex survey data. *Biostatistics* **11**, 4856.
- [24] SHE, D., ZHANG, H., and LI, Z. (2009b). Testing Hardy-Weinberg equilibrium using family data from complex surveys. *Ann Hum Genet* **73**, 449–455.
- [25] SLAGER, S. L. and SCHAID D. J. (2001). Case-control studies of genetic markers: Power and sample size approximations for Armitage’s test for trend. *Hum Hered* **52**, 149–153.
- [26] SPIELMAN, R. S., MCGINNIS, R. E., and EWENS, W. J. (1993). Transmission test for linkage disequilibrium: The insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Ann Hum Genet* **52**, 506–516.
- [27] THORNTON, T. and MCPEEK, M. S. (2010). ROADTRIPS: Case-control association testing with partially or completely unknown population and pedigree structure. *Ann Hum Genet* **86**, 172–184.
- [28] WANG, W., BARRATEE, B., CLAYTON, D., and TODD, J. (2005). Genome-wide association studies: Theoretical and practical concerns. *Nat Rev Genet* **6**, 109–118.
- [29] ZHENG, G., FREIDLIN, B., LI, Z., and GASTWIRTH J. L. (2005). Genomic control for association studies under various genetic models. *Biometrics* **61**, 187–193. [MR2135859](#)
- [30] ZHENG, G. and GASTWIRTH, J. L. (2006). On estimation of the variance in Cochran-Armitage trend tests for genetic association using case-control studies. *Statistics in Medicine* **25**, 3150–3159. [MR2247233](#)

Dewei She

Department of Statistics

The George Washington University

Washington, DC 20052

USA

Hong Zhang

Institute of Biostatistics

School of Life Science

Fudan University

Shanghai 200433

P.R. China

Biostatistics Branch

Division of Cancer Epidemiology and Genetics

National Cancer Institute

Rockville, MD 20852

USA

Yan Li

Department of Mathematics

University of Texas at Arlington

Arlington, TX

USA

Joint Program in Survey Methodology

University of Maryland

College Park, MD

USA

Barry I. Graubard  
Biostatistics Branch  
Division of Cancer Epidemiology and Genetics  
National Cancer Institute  
Rockville, MD 20852  
USA

Zhaohai Li  
Department of Statistics  
The George Washington University  
Washington, DC 20052  
USA

Biostatistics Branch  
Division of Cancer Epidemiology and Genetics  
National Cancer Institute  
Rockville, MD 20852  
USA  
E-mail address: [zli@gwu.edu](mailto:zli@gwu.edu)