

## Adjusting for Treatment Effect when Estimating or Testing Genetic Effect is of Main Interest

Yuanjia Wang<sup>1</sup> and Yixin Fang<sup>2</sup>

<sup>1</sup>Columbia University and <sup>2</sup>Georgia State University

*Abstract:* It is known that “standard methods for estimating the causal effect of a time-varying treatment on the mean of a repeated measures outcome (for example, GEE regression) may be biased when there are time-dependent variables that are simultaneously confounders of the effect of interest and are predicted by previous treatment” (Hernán *et al.* 2002). Inverse-probability of treatment weighted (IPTW) methods are developed in the literature of causal inference. In genetic studies, however, the main interest is to estimate or test the genetic effect rather than the treatment effect. In this work, we describe an IPTW method that provides unbiased estimate for the genetic effect, and discuss how to develop a family-based association test using IPTW for family-based studies. We apply the developed methods to systolic blood pressure data in Framingham Heart Study, where some subjects took antihypertensive treatment during the course of study.

*Key words:* FBAT, hypertension, inverse probability of treatment weighting, systolic blood pressure, time-varying confounding.

### 1. Background

Originating in 1948, Framingham Heart Study (FHS, Dawber *et al.* 1951) is an ongoing prospective study of risk factors including genetic variants for cardiovascular disease. The clinical data from FHS includes blood pressure, total cholesterol, fasting HDL cholesterol, blood glucose and others, and the genetic data includes microsatellite markers and single-nucleotide polymorphisms (SNPs). This study provides opportunities to dissect genetic structure of blood pressure, cholesterol and other complex traits.

Since FHS started, one of the phenotypes, systolic blood pressure (SBP), has drawn many researchers' attention. For example, investigators in Levy *et al.* (2000) and Hunt *et al.* (2002) found genetic markers significantly associated with SBP. Interestingly, as pointed out in Shea *et al.* (1985), some subjects took antihypertensive treatment during the course of study. For these subjects, their

underlying SBPs were distorted by the antihypertensive medication. In Levy *et al.* (2000) and Hunt *et al.* (2002), the SBP levels were adjusted for treatment before analysis by deriving adjusted residuals for subjects on treatment. It is well understood that the past SBP is a time-dependent confounder for estimating the treatment effect because it is a predictor of both the treatment assignment and the future SBP (Robins 1998). With the presence of such time-dependent confounder, standard approaches including adding confounder as a covariate do not work (Robins 1998). With estimating treatment effect as the main interest, Robins and colleagues proposed several remedies including inverse-probability of treatment weighting (IPTW). See, for example, Hernán *et al.* (2002) and Robins (1999).

In genetic studies, the main interest is the genetic effect that would have been observed if every subject was untreated. Consider a gene that increases SBP, therefore carriers are more likely to develop hypertension and more likely to receive antihypertensive medication. Assume that the treatment is effective in reducing subjects' SBPs. Then using the observed SBPs to estimate the genetic effect will create a downward bias because SBPs on carriers who are treated are lower than the underlying true values had they not received treatment. When the genetic effect is of interest, how does one adjust for subjects taking treatment which distorts their SBPs? Does the convenient practice of including confounder as a covariate work?

To answer these questions, understanding the unbalanced roles of the genetic effect and the treatment effect is essential: when the treatment effect is of primary interest, the genotype, which is not affected by the treatment, can be treated as a covariate; when the genetic effect is of primary interest, the treatment, which is predicted by the genotype through the underlying SBP, however, cannot be treated as a covariate. Therefore, including treatment as a covariate to adjust for bias is inappropriate. For cross-sectional studies, authors in Tobin *et al.* (2005) examined the biasedness of several standard methods and proposed to use censored regression model. It is shown in Tobin *et al.* (2005) that by including the treatment assignment as a covariate, some of the genetic effect is adjusted away.

In longitudinal studies, the baseline SBP measurements, based on which the treatment is assigned, may be available. If there are two measurements, one measured when the treatment is assigned and the other measured after the treatment is taken for some period, regular regression analysis with the baseline SBP as a covariate provides an unbiased estimate of the treatment effect, but a biased estimate of the genetic effect. When there are more than two measurements, regular regression analysis gives biased estimates for both the treatment and the genetic effect.

In the next section, we describe an IPTW method that provides asymptoti-

cally unbiased estimate for the genetic effect. By an example and some simulation studies, we quantify the biases in some standard methods and compare them with the proposed methods. We then apply the developed methods to the SBP data in FHS. We conclude with some discussions on how to develop a family-based association test using IPTW for family-based studies.

## 2. An IPTW Method for Estimating the Genetic Effect

We describe an IPTW method for estimating the genetic effect of a candidate gene in longitudinal studies when subjects' phenotypes are distorted by taking treatment. The focus here is the genetic effect in the population where everyone is untreated. Let  $i$  index the subjects and  $j$  index the visits. Let  $Y_{ij}$  denote the phenotype measurement at visit  $j$  of subject  $i$ ,  $i = 1, \dots, n$ , and  $j = 0, 1, \dots, m_i$ . Assume that measurements between subjects are independent, and at the baseline  $j = 0$ , no subject takes the treatment. Let  $Y_i$  be the vector  $(Y_{i1}, \dots, Y_{im_i})^T$ . The predictors include the genotype at the candidate gene, the treatment assignment, and may include time-constant variables such as gender and time-dependent variables such as age.

The standard generalized estimating equation (GEE) regression solves

$$\sum_i \left( \frac{\partial \mu_i(\beta)}{\partial \beta} \right)^T V_i^{-1} (Y_i - \mu_i(\beta)) = 0, \quad (2.1)$$

where  $\mu_i(\beta)$  is the specified marginal means of  $Y_i$  in terms of parameter  $\beta$  and the predictors, and  $V_i$  is the working covariance matrix of  $Y_i$ . For example in a candidate gene study,  $\mu_i(\beta)$  can be parametrized as a linear model,  $X_i\beta$ , where  $X_i$  includes carrier status at the gene and other important covariates such as gender and age to be adjusted. The standard GEE gives a consistent estimate for the genetic effect when there is no time-dependent confounder; see for example Liang and Zeger (1986).

It is shown in Robins (1986) that in the presence of time-varying confounder, the standard GEE may provide biased estimate for the treatment effect even after including past treatment history and other confounders as covariates. When the genetic effect is of interest, in Section 3 we demonstrate that the standard methods provide biased estimates for the genetic effect. We develop an IPTW-type method to find asymptotically unbiased estimate for the genetic effect.

Simply put, the IPTW method solves a weighted GEE with time- and subject-specific weights. The basic idea is that, by properly weighting subjects, we construct a pseudo-population in which the past SBP and the treatment assignment are not associated, and therefore the genetic effect is not distorted by the treatment effect. It is then straightforward to estimate the genetic effect and/or its interaction with environmental effects in the pseudo-population.

To construct IPTW estimators, one can use nonstabilized weight, which is the inverse of the conditional probability of the observed treatment status given the corresponding subject's treatment history and observed time-dependent confounder. As an alternative, stabilized weight proposed in [6] is shown to be more efficient and stable, especially when the conditional probability is small. For subject  $i$ , the stabilized weights are defined as,

$$SW_i(0) = 1, \quad SW_i(j) = \prod_{k=1}^j \frac{Pr(T_i(k)|T_i(k-1), Z_i)}{Pr(T_i(k)|\bar{T}_i(k-1), \bar{C}_i(k-1))}, \quad j = 1, \dots, m_i, \quad (2.2)$$

where  $Z_i$  comprises the baseline covariates believed to predict the treatment assignment,  $T_i(k)$  is the indicator of taking treatment at visit  $k$ ,  $\bar{T}_i(k)$  is the treatment history up to visit  $k$ ,  $C_i(k)$  is the vector of time-varying variables including the confounder at visit  $k$ , and  $\bar{C}_i(k)$  is the history of  $C_i(k)$  up to visit  $k$ . Let  $SW_i$  denote the matrix with diagonal elements  $SW_i(1), \dots, SW_i(m_i)$ . The IPTW method solves the weighted estimating equation,

$$\sum_i \left( \frac{\partial \mu_i(\beta)}{\partial \beta} \right)^T V_i^{-1} SW_i (Y_i - \mu_i(\beta)) = 0. \quad (2.3)$$

In practice, the weights (2.2) are unknown, but they can be estimated from the data by specifying a model (e.g., logistic model) for the treatment status  $T_i(k)$ ; see [1] for more discussion. We illustrate how to obtain the weights in the next section.

### 3. Example and Simulations

#### 3.1 Example

We assume there are 200 subjects, half with genotype  $G = 1$  and the other half with  $G = 0$ . The threshold for defining high SBP is 130 mmHg. Let  $L$  indicate high SBP, with  $L = 1$  being having high SBP and  $L = 0$  being having low SBP. The probability of developing high SBP,  $P(L = 1|G)$ , is higher in subjects with  $G = 1$  than those with  $G = 0$ ; they are 0.2 and 0.1, respectively. The assumptions and results are summarized in Table 1. Here we assume that the numbers of subjects within each categories of  $L$  and  $G$  are consistent perfectly with their expectations, as shown in the 6th column of Table 1. At time 1, the mean SBP of subjects with  $L = 1$  is 140 mmHg, and those with  $L = 0$  is 120 mmHg. Their SBPs are recorded before some subjects being prescribed and then starting to take treatment. Assume that among subjects with  $L = 1$ , 80% will receive treatment, and among those with  $L = 0$ , 10% will receive treatment. Assume the treatment reduces SBP by an average of 20 mmHg at time 2.

Table 1: An example illustrating biases in various methods

$G$	$L$	$T$	$P(L G)$	$P(T L)$	$Num$	$E(Y_1 G, L)$	$E(Y_2 G, L, T)$	$W$	$N_w$	$SW$	$N_{sw}$
1	1	1	0.2	0.8	16	140	120	1.25	10	0.26	4.1
1	1	0	0.2	0.2	4	140	140	5.00	10	3.98	15.9
1	0	1	0.8	0.1	8	120	100	10.00	40	2.05	16.4
1	0	0	0.8	0.9	72	120	120	1.11	40	0.88	63.6
0	1	1	0.1	0.8	8	140	120	1.25	5	0.26	2.05
0	1	0	0.1	0.2	2	140	140	5.00	5	3.98	7.95
0	0	1	0.9	0.1	9	120	100	10.00	45	2.05	18.45
0	0	0	0.9	0.9	81	120	120	1.11	45	0.88	71.55

For simplicity, no disturbance is added to the phenotype. The outcomes therefore coincide with their expectations. The expected bias is the same as shown here if adding a mean zero random disturbance to the SBP. The genetic effect is

$$\beta = [20(140) + 80(120)]/100 - [10(140) + 90(120)]/100 = 2.$$

If only using data at time 1, we can obtain an unbiased estimate of the genetic effect, but in practice it is more advantageous to use all the available data.

We consider four standard methods: (1) ignore that some subjects have taken treatments; (2) exclude subjects taking treatments; (3) treat  $T$  as a covariate; (4) treat both  $L$  and  $T$  as covariates. For method 1, the estimated genetic effect is

$$\begin{aligned} & [140(16 + 4) + 120(8 + 72) + 120(16) + 140(4) + 100(8) + 120(72)]/200 \\ & - [140(8 + 2) + 120(9 + 81) + 120(8) + 140(2) + 100(9) + 120(81)]/200 = 1.3. \end{aligned}$$

Therefore method 1 underestimates the genetic effect by 0.7. Similar calculations show that method 2 and 3 underestimate the genetic effect by 0.58 and 0.33, respectively. Method 4 estimates the genetic effect as zero and the treatment effect as  $-20$ . Therefore, by adjusting for the baseline SBP and treatment assignment, the estimate of the treatment effect is unbiased but the estimate of the genetic effect is biased. Treating  $T$  as covariate has the smallest bias, but still non-negligible.

The IPTW method requires computing weights. At time 1, the weights are one. At time 2, the unstabilized weights are  $1/Pr(T|L)$ , and the stabilized weights are  $Pr(T)/Pr(T|L)$ . For instance, the stabilized weight in the first row of Table 1 is  $(16 + 8 + 8 + 9)/200/0.8 = 0.26$ , and in the second row is  $(1 - (16 + 8 + 8 + 9)/200)/0.2 = 3.98$ . These weights are summarized in the 11th column of Table 1. The weighted sample sizes  $N_w$  and  $N_{sw}$  are obtained by multiplying the

original sample sizes by the corresponding weights and then normalize so that the total sample size remains as 200. These are summarized in the 10th and the 12th column.

The IPTW estimators can now be computed using the adjusted sample sizes. Using the stabilized weights, the estimate of the genetic effect is

$$\begin{aligned} & [a_1(16 + 4) + a_2(8 + 72) + a_3(4.1) + a_4(15.9) + a_5(16.4) + a_6(63.6)]/200 \\ & - [a_1(8 + 2) + a_2(9 + 81) + a_3(2.05) + a_4(7.95) + a_5(18.45) + a_6(71.55)]/200 \\ & = 2, \end{aligned}$$

where  $a_1 = 140, a_2 = 120, a_3 = 120, a_4 = 140, a_5 = 100, a_6 = 120$ . Similarly, using the nonstabilized weights the estimate is also 2. Although both the stabilized and nonstabilized weights give unbiased estimates, the stabilized weights generally provide more efficient and stable estimator Robins (1986). It can be seen in the example that the range of the nonstabilized weights is wider than the stabilized weights.

### 3.2 Simulations

Simulation studies were conducted based on a modification of Example 1. Two hundred subjects were simulated. Genotypes were generated with  $P(G = 1) = 0.4$ . The effect of having  $G = 1$  was assumed to increase a subject's probability of developing hypertension. Four scenarios were considered: subjects with  $G = 1$  were simulated to be hypertensive ( $L = 1$ ) at time 1 with probability 0.2, 0.3, 0.4 or 0.5, and subjects with  $G = 0$  with probability 0.1. This is equivalent to that subjects with  $G = 1$  on average had SBPs 2, 4, 6, or 8 mmHg higher than those with  $G = 0$ , provided that  $E(Y_1|G, L)$  followed the 7th column of Table 1.

Treatments were assigned to subjects following probabilities  $P(T|L)$  in the 5th column of Table 1. Hypertensive subjects had higher probability to be treated. To account for a natural increasing trend of SBP as subjects get older, each subject was assumed to have underlying SBP 2 mmHg higher at time 2 than time 1. The treatment effect was generated from a normal distribution with mean -22 and standard deviation 3. The random error added to the underlying SBP at the two time points were generated from a bivariate normal distribution with mean 0, standard deviation 10 for both components, and correlation coefficient 0.5. Each experiment was repeated 1000 times.

## 4. Results

### 4.1 Simulation results

Table 2 summarizes the results of estimating and testing the genetic effect. In all scenarios, the standard approaches (no adjustment, excluding subjects on treatment, and including treatment indicator as a covariate) underestimate the genetic effect. The bias increases with the strength of the genetic effect. Including treatment as a covariate has smaller bias than the other two approaches, but the bias is still non-ignorable. Both IPTW approaches with nonstabilized and stabilized weights provide estimates with minimal bias (absolute value ranging from 0.01 to 0.07), but the standard errors using stabilized weights are smaller than using nonstabilized weights. Comparing with the standard approaches, the IPTW approaches have larger standard errors. This reflects a tradeoff between the bias and the variance.

Table 2: Estimation of genetic effect ( $\beta$ ) in four scenarios

Scenario	Approach	Mean $\hat{\beta}$	Bias	Emp. SE*	Est. SE <sup>†</sup>	Power	Power Loss <sup>‡</sup>
$\beta = 2$	No adjustment	1.22	-0.78	1.7	1.6	0.13	27.8%
	Exclude being treated	1.38	-0.62	1.7	1.6	0.16	11.1%
	Treatment as covariate	1.63	-0.37	1.7	1.7	0.17	5.6%
	Non-stabilized IPTW	2.01	0.01	2.3	2.2	0.16	11.1%
	Stabilized IPTW	1.99	-0.02	2.1	2.0	0.18	-
$\beta = 4$	No adjustment	2.49	-1.51	1.7	1.7	0.33	36.5%
	Exclude being treated	2.84	-1.16	1.8	1.8	0.37	28.8%
	Treatment as covariate	3.26	-0.74	1.8	1.8	0.47	9.6%
	Non-stabilized IPTW	4.1	0.1	2.4	2.3	0.45	13.5%
	Stabilized IPTW	4.05	0.05	2.2	2.1	0.52	-
$\beta = 6$	No adjustment	3.67	-2.33	1.9	1.8	0.51	30.1%
	Exclude being treated	4.45	-1.55	2.0	2.0	0.63	13.7%
	Treatment as covariate	4.84	-1.16	2.0	1.9	0.70	4.3%
	Non-stabilized IPTW	6.02	0.02	2.7	2.5	0.64	15.1%
	Stabilized IPTW	5.96	-0.04	2.3	2.3	0.73	-
$\beta = 8$	No adjustment	4.86	-3.14	1.8	1.7	0.80	16.7%
	Exclude being treated	6.04	-1.96	1.9	1.8	0.89	7.3%
	Treatment as covariate	6.49	-1.51	1.9	1.8	0.93	3.1%
	Non-stabilized IPTW	8.07	0.07	2.5	2.3	0.91	5.2%
	Stabilized IPTW	7.98	-0.02	2.2	2.1	0.96	-

Note:

\*: Empirical Standard Error, \* Empirical standard error, <sup>†</sup> Mean estimated standard error, <sup>‡</sup> Percentage decrease of power comparing with IPTW approach using stabilized weights.

Table 2 also shows the results of testing for genetic effect using Wald statistic based on sandwich variance estimate. The IPTW approach with stabilized weights has the largest power. Comparing to stabilized IPTW, the percentage

loss of power of the other approaches ranges from 3% to 37%. The most popular standard approach of adjusting for treatment effect by including it as a covariate can have a power loss of 9.6% (scenario 2).

For estimating treatment effect, including treatment assignment as a covariate has large bias ranging from 8.95 to 9.47 (results not shown). The IPTW approach has negligible bias ranging from 0.01 to 0.15. The estimator of treatment effect behaves similarly across all scenarios regardless of the size of the genetic effect.

## 4.2 Data analysis results

We apply the proposed methods to analyze the SBP data in FHS. There were 5209 subjects recruited (2336 men and 2873 women) from Framingham, Massachusetts. FHS includes longitudinal observations from two cohorts: Cohort 1, the original Framingham cohort, was first examined in 1948 and had been examined every 2 years thereafter; Cohort 2, composed primarily of offspring of the original cohort and the spouses of these offspring, was examined first in 1971 and had been examined approximately every 4 years Dawber *et al.* (1951).

In the mid-1990s about 1800 members of the largest 330 pedigrees in FHS were selected for a genome scan conducted by the Mammalian Genotyping Service in Marshfield, Wisconsin (Cupples *et al.* 2003). The genetic markers span the whole genome with a 10-cM density. Researchers in Levy *et al.* (2003) identified a marker, GATA25A04 (D17S1299) on chromosome 17 with a two point LOD score of 3.8 for the phenotype SBP. We focus on analyzing the genetic effect at this marker. We estimated the genetic effect using IPTW method and compared it with the standard methods.

We fitted three standard GEE regressions with marker alleles, age, sex and BMI as covariates: the first ignored that some subjects were treated by anti-hypertensive treatment; the second excluded treated subjects; and the third added the treatment as a covariate in addition to the original ones. We also applied the IPTW method, where the weights  $SW_i$  were computed using (2.2). Here  $Z_i$  includes sex and BMI,  $C_i(k)$  includes the previously observed SBPs and age. The conditional probabilities in the numerator and the denominator of the weights were fitted by logistic regression models; for the numerator, the treatment assignment was outcome and sex and BMI were predictors; for the denominator, past SBPs, age, and age squared were added as predictors.

For all methods, each marker allele was coded as a binary variable of having at least one copy of the allele or not. We assumed that the observations from different families were independent and used independent working correlation matrix. Among 1701 subjects being genotyped, 1667 have valid data at the marker GATA25A04. The marker has eight alleles, with frequencies ranging from 0.1% to 38%. We combined the two alleles with least frequencies (alleles



184 and 212).

Table 3: Estimating genetic effect of maker GATA25A04

Variable	No Adjustment			Exclude Being Treated			Treatment as Covariate			Stabilized IPTW		
	Est.	SE	<i>p</i> value	Est.	SE	<i>p</i> value	Est.	SE	<i>p</i> value	Est.	SE	<i>p</i> value
x184212 (1%)*	-7.71	5.33	0.15	-7.94	4.86	0.10	-7.27	4.85	0.13	-7.15	4.82	0.14
x188 (4%)	1.88	1.33	0.16	1.95	1.29	0.13	2.05	1.29	0.11	2.40	1.43	0.09
x192 (12%)	0.85	0.93	0.36	0.78	0.91	0.39	0.85	0.90	0.35	1.47	0.97	0.13
x196 (45%)	1.69	1.07	0.12	1.63	1.04	0.12	1.74	1.03	0.09	3.40	1.34	0.01
x200 (29%)	1.33	0.91	0.15	1.33	0.90	0.14	1.32	0.89	0.14	2.52	1.26	0.04
x204 (6%)	2.71	1.21	0.03	2.66	1.23	0.03	2.71	1.19	0.02	3.93	1.70	0.02
x208 (3%)	1.48	1.59	0.35	1.14	1.60	0.48	1.39	1.57	0.38	2.27	1.66	0.17

\* Allele frequency

Table 3 reports the results from the four methods. We found that the estimate of the genetic effect from the IPTW method was larger than those from the standard methods, except for the combination of alleles 184 and 212 where the frequency is 1%. The alleles 196 and 200 were not significant in the standard methods, but were significant in the IPTW method; for example, the allele 196 had an estimate of 3.40 ( $p = 0.01$ ) in the IPTW method, but the standard methods failed to detect it. We also found that the allele 204 had a significant effect in all methods, but the IPTW provided the largest estimate. These results are consistent with the finding that the standard methods underestimate the genetic effect.

## 5. Discussions

Here we describe an IPTW-type method to provide asymptotically unbiased estimators for the genetic effect when the genetic effect is confounded by the treatment effect. By constructing proper weights, we break the correlation between the treatment and the past SBP and therefore the correlation between the treatment and genetic factors because genetic factors predict treatment assignment through SBP. For the method to be valid, the assumption of no unmeasured confounder is assumed. This assumption is debatable, but by adjusting for the observed confounders, the proposed methods can improve the standard methods such as GEE regression to some extent. For the SBP data in FHS, the assumption of no unmeasured confounder is that the treatment assignment depends completely on the measured past SBPs and other covariates. This assumption will not hold if the subjects in the study have their own physicians outside of the study who also monitor their SBPs and prescribe the antihypertensive treatment and the measurements taken outside of FHS are not available to the FHS investigators. However, since subjects are examined every two years in FHS, the SBP measured

by an outside physician could be close to the one measured at the nearest preceding visit by the FHS researchers. In this case, it is still reasonable to assume that there are no unmeasured confounder.

Here we assume that the subjects are not taking the treatment at the beginning of the study. While this is true for Framingham study because it started in an era of no available antihypertensive treatment, in other applications this assumption may not hold. If some of the subjects are treated at the baseline because they already have high blood pressure and if their SBP measurements based on which the treatment is prescribed are not available, the weights  $SW_i$  cannot be computed. These subjects will be excluded from the analysis. However, if the blood pressure measurements used to prescribe the treatments before the study begins are available, these subjects can still contribute to the analyses.

Using the IPTW estimator in (2.3) and its sandwich variance estimate, it is straightforward to construct a Wald test for the genetic effect. For multivariate or longitudinal data in family studies, to control for population stratification, a family-based association test using GEE regression (FBAT-GEE) is proposed in Lange *et al.* (2003). The test statistic is

$$\frac{(S - E(S|\mathcal{C}))^2}{V(S|\mathcal{C})}, \quad (5.1)$$

where  $S = \sum_i G_i \Delta_i V_i^{-1} Y_i$ ,  $G_i$  is the candidate gene's genotype, and  $G_i \Delta_i$  is the slope of  $\mu_i(\beta)$  at the null hypothesis of no genetic effect, with  $\mu_i(\beta)$  and  $V_i$  defined in (2.1). The expectation  $E(S|\mathcal{C})$  and variance  $V(S|\mathcal{C})$  are computed conditionally on the minimal sufficient statistic  $\mathcal{C}$  under the null hypothesis; the minimal sufficient statistics are discussed in Rabinowitz and Laird (2000). Under the null hypothesis, the test statistic has a chi-square distribution with degrees of freedom equal to the number of the levels of  $G_i$  minus one.

Although the FBAT approach has correct type-I error using any test statistic  $S$ , a carefully chosen statistic can be more powerful. For the SBP data, adjusting for the influence of antihypertensive treatment may improve testing power. To this aim, one can use an FBAT-type test based on IPTW (FBAT-IPTW). The test statistic is

$$\frac{(T - E(T|\mathcal{C}))^2}{V(T|\mathcal{C})}, \quad (5.2)$$

where  $T = \sum_i G_i \Delta_i V_i^{-1} SW_i Y_i$ , with  $G_i$ ,  $\Delta_i$  and  $V_i$  defined in (5.1) and  $SW_i$  defined in (2.2). The expectation  $E(T|\mathcal{C})$  and variance  $V(T|\mathcal{C})$  are also computed conditionally on the minimal sufficient statistic  $\mathcal{C}$  under the null hypothesis. Following the arguments in Lange *et al.* (2003) and Rabinowitz and Laird (2000), we see that the FBAT-IPTW is robust to population stratification. Note that

because observed phenotypes are a part of  $\mathcal{C}$  and the weights  $SW$  only involves observed phenotypes and covariates other than genotypes, the weights can be taken out of the conditional expectation. Therefore in practice, the FBAT-IPTW can be calculated using the software FBAT (<http://www.biostat.harvard.edu/~fbat/>) by first computing a weighted phenotype  $SW_i Y_i$  and then inputting them as phenotype in the FBAT program.

### Acknowledgement

The data was obtained from the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. This work was supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No. N01-HC-25195). Yuanjia Wang (PI) and Yixin Fang (Subcontract PI)'s research is supported by NIH AG031113-01A2.

### References

- Cupples, L. A., Yang, Q., Demissie, S., Copenhaver, D. and Levy, D. (2003). Description of the Framingham Heart Study data for Genetic Analysis Workshop 13. *BMC Genet* **4**(suppl 1), S2.
- Dawber, T. R., Meadors, G. F. and Moore, F. E. (1951). Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health* **41**, 279-286.
- Hernán, M. A., Brumback, B. A. and Robins, J. M. (2002). Estimating the casual effect of zidovudine on CD4 count with a marginal structural model for repeated measures. *Stat Med* **21**, 1689-1709.
- Hunt, S. C., Ellison, R. C., Atwood, L. D., Pankow, J. S., Province, M. A. and Leppert, M. F. (2002). Genome scans for blood pressure and hypertension: the National Heart, Lung, and Blood Institute family heart study. *Hypertension* **40**, 1-6.
- Lange, C., Silverman, E. K., Xu, X., Weiss, S. T. and Laird, N. M. (2003). A multivariate family-based association test using generalized estimating equations: FBAT-GEE. *Biostatistics* **4**, 195-206.
- Levy, D., DeStefano, A. L., Larson, M. G., O'Donnell, C. J., Lifton, R. P., Gavras, H., Cupples, L. A. and Myers, R. H. (2000). Evidence for a gene influencing blood pressure on chromosome 17. Genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the Framingham Heart Study. *Hypertension* **36**(4), 477-483.
- Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13-22.

- Rabinowitz, D., Laird, N. M. (2000). A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. *Human Hered* **50**, 211-223.
- Robins, J. M. (1986), A new approach to causal inference in mortality studies with a sustained exposure period application to the healthy worker survivor effect. *Mathematical Modeling* **7**, 1393-1512.
- Robins, J. M. (1998). Marginal structural models. In *1997 Proceedings of the Section on Bayesian Statistical Science*. American Statistical Association: Alexandria, VA. 1-10.
- Robins, J. M. (1999). Marginal structural models versus structural nested models as tools for causal inference. In *Statistical Models in Epidemiology: The Environmental and Clinical Trials* (Edited by Halloran E, Berry D., 95-134.) Springer-Verlag.
- Shea, S., Cook, E. F., Kannel, W. B. and Goldman, L. (1985). Treatment of hypertension and its effect on cardiovascular risk factors: data from the Framingham Heart Study. *Circulation* **71**, 22-30.
- Tobin, M. D., Sheehan, N. A., Scurrah, K. J. and Burton, P. R. (2005). Adjusting for treatment effects in studies: antihypertensive therapy and systolic blood pressure. *Stat. Med.* **24**, 2911-2935.

Received April 30, 2009; accepted August 26, 2009.

Yuanjia Wang  
Department of Biostatistics  
Mailman School of Public Health  
Columbia University  
722 W. 168th St.  
New York, NY 10032, USA  
yw2016@columbia.edu

Yixin Fang  
Department of Mathematics and Statistics  
Georgia State University  
750 COE, 7th floor, 30 Pryor Street  
Atlanta, GA 30303, USA  
matyxf@langate.gsu.edu