Statistica Sinica **22** (2012), 337-357
doi:http://dx.doi.org/10.5705/ss.2009.303

A PATTERN-MIXTURE MODEL FOR HAPLOTYPE ANALYSIS OF LONGITUDINAL TRAITS WITH NON-IGNORABLE DROPOUT

Hongying Li and Rongling Wu

University of Florida and Pennsylvania State University

Abstract: Current statistical methods allow the characterization of DNA sequence variants associated with interpersonal differences in a complex biological response. However, this process is significantly hindered when some subjects have to drop out early due to physiological side effects or limited duration. Here, we derive a pattern-mixture model for detecting functional nucleotide combinations (or haplotypes) responsible for longitudinal responses by making full use of information from those dropout data. The model was formulated within the maximum likelihood context, with the model parameters, haplotype frequencies, and haplotype effects estimated by implementing the EM and Newton-Raphson algorithms. One advantage of the model is to generate and address a number of clinically meaningful hypotheses about the genetic control mechanisms of longitudinal responses and time-to-event processes. By analyzing a pharmacogenomic data set, the model identified significant haplotype effects on heart rate increases in response to increasing doses of dobutamine. The statistical properties of the model and its usefulness and utilization were investigated through computer simulation. The new model can be used to unravel the genetic architecture of interpersonal variation in complex longitudinal responses with incomplete data and ultimately to materialize the idea of clinical genomics.

Key words and phrases: Drug response, EM algorithm, functional mapping, longitudinal trait, non-ignorable dropout, pattern mixture, quantitative trait nucleotides.

1. Introduction

Variations in nucleotides at particular locations are called single nucleotide polymorphisms (SNPs). An example of a SNP is the difference of the DNA segment AAGGTTA for individual 1 from ATGGTTA for individual 2, where the second nucleotides from the left end form a polymorphism A/T. The linear combination of alleles at different SNPs that are transmitted together on the same chromosomal region is called the haplotype. Many studies, through statistical simulation, show that haplotypes composed of multiple SNPs could better explain variation in a phenotypic trait than single SNPs (Collins, Guyer, and Charkravarti (1997); Akey, Jin, and Xiong (2001); Morris and Kaplan (2002);

Zaykin et al. (2002)). There has also been a vast body of molecular evidence indicating that haplotypes are associated with many aspects of drug response (Judson, Stephens, and Windemuth (2000); Bader (2001); Rha et al. (2007)). However, a direct analysis of association between haplotypes and phenotypes may be difficult because there is currently no easy way to genotype haplotypes. For this reason, powerful statistical models for a missing data problem have been derived to estimate genetic effects and variation due to haplotypes (Liu et al. (2004); Lin and Zeng (2006); Lin and Huang (2007); Huang, Amos, and Lin (2007)).

When haplotype analysis is used to study the genetic control of drug response, we face two substantial issues characterized by this trait. First, drug response presents a dynamic process in which the efficacy and toxicity of a medication are functions of drug concentration and time. Second, some patients in clinical trials drop from the study because of physiological or other unavoidable reasons. Although a conceptual model, called functional mapping, has been derived to study the genetic architecture of dynamic traits (Ma, Casella, and Wu (2002); Wu and Lin (2006)), it is still unclear how early dropouts affect our statistical inference about genetic control. We address the second issue.

To better describe our problem, we start with a pharmacogenetic study. A group of 163 patients participated in a study in which all people were measured for heart rate repeatedly after treatment of dobutamine (Figure 1). The patients received increasing doses of dobutamine, until they achieved a target heart rate response or predetermined maximum dose. Of these subjects studied, 112 (69%) completed the tests of heart rate at all the six dose levels; the others dropped out before the completion of the trial because heart rates at any higher dose level were beyond their physiological limits. A total of 31 (19%), 15 (9%), and 5 (3%) subjects dropped out after receiving four, three, and two injections, respectively. Because the dropouts of these subjects were likely related to the outcome, they are called non-ignorable dropouts. Existing models for functional mapping are not sufficient to analyze non-ignorable dropouts.

An effective analysis of non-ignorable dropout data is based on pattern-mixture models that construct a likelihood on the joint distribution of the complete response and dropout mechanism, and factors the joint likelihood as the marginal distribution of the mechanism multiplied by the conditional distribution of the response given the mechanism (Wu and Bailey (1989); Little (1993, 1995)). Pattern-mixture models have now been used in many applications for which longitudinal non-ignorable missing data are common (Fitzmaurice, Laird, and Shneyer (2001); Hogan and Laird (1997)). We integrate pattern-mixture models within the framework of functional mapping through explicit modeling of the missing data distribution by first identifying different patterns of missing

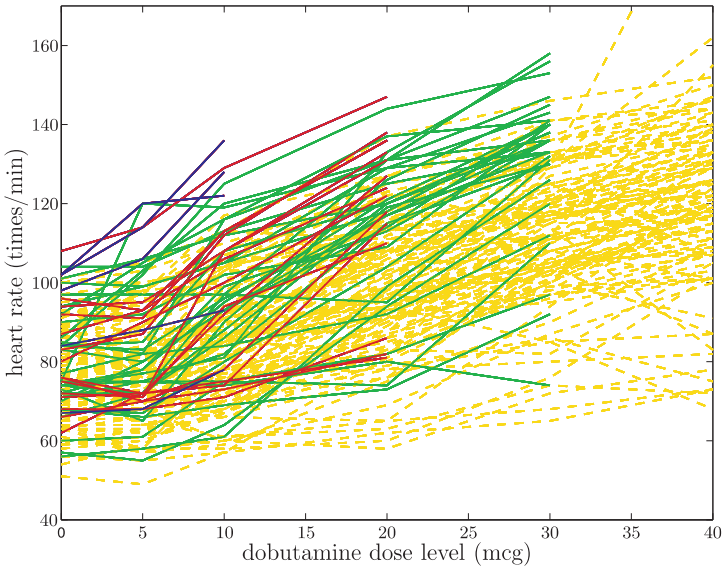

Figure 1. The dobutamine drug response experiment data. The curves in dashed lines are the complete responses covering all dose levels, whereas the curves in solid lines are the incomplete responses showing nonignorable dropouts from different dosages.

data and then including parameters in the outcomes model that capture this effect. Parameter estimation is implemented with the EM and Newton-Raphson algorithm. We formulate a number of hypotheses tests regarding the genetic control of longitudinal responses and dropout times. Simulation studies and data analyses were used to demonstrate the power and usefulness of the model.

2. Model

2.1. Genetic and clinical designs

Suppose a random sample of size N is drawn from a natural human population that is assumed to be at Hardy-Weinberg equilibrium (HWE). For an illustration of our model, we consider a simple case in which a haplotype is composed of two SNPs, **A** (with alleles A and a) and **B** (with alleles B and b). The capital alleles are symbolized as 1 and the small alleles as 0. There is no technical difficulty in extending the model to study the effects of haplotypes containing an arbitrary number of SNPs. The two SNPs considered form four haplotypes $[AB]$, $[Ab]$, $[aB]$, and $[ab]$, with the haplotype frequencies designated as p_{11}, p_{10}, p_{01}, and p_{00}, respectively. The four haplotypes derived from the maternal ($\mathtt{m}$) and paternal parents ($\mathtt{p}$) unite randomly to generate 10 diplotypes $[AB][AB]$, $[AB][Ab]$, $[AB][aB]$, $[AB][ab]$, $[Ab][Ab]$, $[Ab][aB]$, $[Ab][ab]$ $[aB][aB]$, $[aB][ab]$, and

$[ab][ab]$. For an HWE population, the frequency of a diplotype is expressed as the product of the frequencies of the two haplotypes that constitute the diplotype.

In most practical studies, only genotypes (i.e., combination between alleles at individual SNPs) are observed because it can be expensive to observe diplotypes. For the double heterozygote Aa/Bb, there are two possible diplotypes $[AB][ab]$ and $[Ab][aB]$ that are not directly observable. When a specific haplotype or diplotype causes variation in drug response, we need to develop a mixture model for inferencing this effect based on observed genotype data. Let $H = [r_1^{\mathtt{m}} r_2^{\mathtt{m}}][r_1^{\mathtt{p}} r_2^{\mathtt{p}}]$ $(r_1^{\mathtt{m}}, r_2^{\mathtt{m}}; r_1^{\mathtt{p}}, r_2^{\mathtt{p}} = 1, 0)$ and $G = r_1 r_1'/r_2 r_2'$ $(r_1 \geq r_1',\ r_2 \geq r_2' = 1,\ 0)$ be a general two-SNP diplotype and genotype, respectively. For subject i, we use H_i and G_i to denote its diplotype and genotype. Genotype G_i follows a multinomial distribution with nine possible genotypes of size $n_{r_1 r_1'/r_2 r_2'}$ with probabilities expressed as the product of the frequencies of the two underlying haplotypes. The likelihood function is as

$$L(\mathbf{\Omega}_p) = (p_{11}^2)^{n_{11/11}} (2p_{11}p_{10})^{n_{11/10}} (p_{10}^2)^{n_{11/00}} (2p_{11}p_{01})^{n_{10/11}} (2p_{11}p_{00} + 2p_{10}p_{01})^{n_{10/10}} (2p_{10}p_{00})^{n_{10/00}} (p_{01}^2)^{n_{00/11}} (2p_{01}p_{00})^{n_{00/10}} (p_{00}^2)^{n_{00/00}}, \tag{2.1}$$

from which the EM algorithm can be employed to obtain the maximum likelihood estimates (MLEs) of unknown vector $\mathbf{\Omega}_p = (p_{11}, p_{10}, p_{01}, p_{00})$ (Liu et al. (2004)).

In a pharmacogenetic study, the subjects genotyped are measured for a pharmacokinetic or pharacodynamic parameter of drug response, repeatedly at multiple time points or dose levels. Let us use the example shown in Figure 1 to describe such a longitudinal trial. Let $\mathbf{y}_i = (y_{i1}, \ldots, y_{iT_i})$ be the vector of heart rates for subject i measured at T_i doses; we use $\mathbf{t}_i = (t_{i1}, \ldots, t_{iT_i})$ to express subject-specific dose levels. Let D_i denote the pre-specified dose at which subject i drops out based on this subject's physiological limit. Given its uncertainty, the maximum dose (C_i) this subject can tolerate is right censored relative to his/her dropout dose (D_i). The indicator of censoring is denoted by $\Delta_i = I(D_i \leq C_i)$, zero if D_i is censored and one otherwise. The observed dropout dose is expressed as $\tilde{D}_i = \min(D_i, C_i)$, where $\tilde{D}_i$ has possible values from $S = \{s_1, \ldots, s_L\}$. Thus, the data for subject i consist of longitudinal measures of heart rate and dropout doses as well as SNP genotypes, denoted as $(\mathbf{y}_i, \mathbf{t}_i, \tilde{D}_i, \Delta_i, G_i)$.

2.2. Haplotyping a longitudinal trait

Assume that there is a risk haplotype from a pool of haplotypes. This risk haplotype (R) has a different genetic value for the longitudinal trait studied from the remaining haplotypes, collectively called the non-risk haplotype $(\bar{R})$ (Liu et al. (2004)). The risk and non-risk haplotypes yield three possible composite diplotypes, RR (coded as 2), $R\bar{R}$ (coded as 1), and $\bar{R}\bar{R}$ (coded as 0). Let Q_i

denote the composite diplotype of subject i, thus having $Q_i = j$ $(j = 2, 1, 0)$. According to quantitative genetic theory (Lynch and Walsh (1998)), we define the vectors of genotypic values for the three composite diplotypes at different doses as

$$\begin{aligned} \mathbf{u}_2 &= \mathbf{u} + \mathbf{a}, \\ \mathbf{u}_1 &= \mathbf{u} + \mathbf{d}, \\ \mathbf{u}_0 &= \mathbf{u} - \mathbf{a}, \end{aligned} \tag{2.2}$$

where $\mathbf{u}$ is the mean vector, $\mathbf{a}$ is the vector of the dose-dependent additive genetic effects due to the substitution of a risk haplotype, and $\mathbf{d}$ is the vector of the dose-dependent dominant genetic effects due to the interaction between the risk and non-risk haplotypes.

The relationship between the observed heart rate vector and genotypic mean vector of a composite diplotype is described by a regression model,

$$\mathbf{y}_i = \sum_{j=0}^{2} \xi_{ij} \mathbf{u}_{j|i} + \mathbf{e}_i, \tag{2.3}$$

where ξ_{ij} is 1 if a composite diplotype j is considered for subject i and 0 otherwise, $\mathbf{u}_{j|i}$ is the genotypic mean vector for subject i who carries composite diplotype j, and $\mathbf{e}_i$ is the residual error vector (i.e., the accumulative effect of polygenes and errors) that is independently and identically distributed normal with mean vector zero and covariance matrix $\mathbf{\Sigma}_i$.

In a pharmacodynamic study, we often use a mathematical equation to describe drug response. For a specific composite diplotype, drug response at dose $t_{i\tau}$ is expressed as

$$u_{j|i}(t_{i\tau}) = g(t_{i\tau}; \mathbf{\Theta}_j), \tag{2.4}$$

where $\mathbf{\Theta}_j$ is the parameters that describe the mathematical equation.

A number of approaches can be used to model the covariance structure of the measurement process. We use the order one structured antedependence (SAD) model for the covariance function (Zimmerman and Núñez-Antón (2001)), in which the dose-dependent variance and covariance are described as

$$\begin{aligned} \mathtt{var}(y_i(t_{i\tau})) &= \frac{1 - \phi^{2t_{i\tau}}}{1 - \phi^2}\sigma^2, \\ \mathtt{cov}(y_i(t_{i\tau_1}), y_i(t_{i\tau_2})) &= \phi^{t_{i\tau_2} - t_{i\tau_1}} \frac{1 - \phi^{2t_{i\tau_1}}}{1 - \phi^2}\sigma^2, \text{ for } t_{i\tau_2} > t_{i\tau_1}. \end{aligned} \tag{2.5}$$

Thus only two parameters, σ^2 and ϕ, are used to model the covariance structure.

2.3. Conditional distributions

The joint density distribution of $(\mathbf{y}_i, D_i, G_i)$ given H_i and Q_i is expressed as

$$\begin{aligned} f(\mathbf{y}_i, D_i, G_i|H_i, Q_i) &= f(G_i|H_i)f(\mathbf{y}_i, D_i|Q_i) \\ &= f(G_i|H_i)f(\mathbf{y}_i|D_i, Q_i)f(D_i|Q_i), \end{aligned} \tag{2.6}$$

where $f(G_i|H_i)$ is the multinomial distribution of SNP genotypes from which to construct the likelihood (2.1), $f(\mathbf{y}_i|D_i, Q_i)$ is the conditional distribution of longitudinal observations given D_i and Q_i, and $f(D_i|Q_i)$ is the conditional distribution of dropout dose given the composite diplotype.

We assume that $f(\mathbf{y}_i|D_i, Q_i)$ is multivariate normal with mean vector $\mathbf{u}_{(j|i)l} = \{g(t_{i\tau}; \mathbf{\Theta}_{jl})\}_{\tau=1}^{T_i}$, specific to subject i with composite diplotype j dropping out at s_l, modeled by a set of curve parameters $\mathbf{\Theta}_{jl}$, and covariance matrix specified by the SAD model (2.5). We also assume that $f(D_i|Q_i)$ is multinominal with possible outcomes from $S = \{s_1, \ldots, s_L\}$,

$$f(D_i|Q_i) = \prod_{j=0}^{2}\prod_{l=1}^{L} \pi_{jl}^{\zeta_{ij}\delta_{il}}, \tag{2.7}$$

where $\zeta_{ij} = I(Q_i = j)$ and $\delta_{il} = I(D_i = s_l)$, $\pi_{jl} = Pr(D_i = s_l|Q_i = j)$, and $\sum_{l=1}^{L} \pi_{jl} = 1$.

2.4. Complete data likelihood

Let $\mathbf{\Omega}_q = (\{\mathbf{\Theta}_{jl}\}_{j=0,l=1}^{2,L}, \rho, \sigma^2)$ denote the quantitative genetic parameter vector related to haplotype effects and residual (co)variance. The likelihood of the complete data $(\mathbf{y}_i, D_i, Q_i)$ is

$$\begin{aligned} L^f(\mathbf{\Omega}_q) &= \prod_{i=1}^{N} L^f(\mathbf{y}_i|D_i, Q_i)L^f(D_i|Q_i)L^f(Q_i) \\ &= \prod_{i=1}^{N}\prod_{l=1}^{L}\prod_{j=0}^{2}\left[(2\pi)^{-T_i/2}|\mathbf{\Sigma}_i|^{-1/2}\exp\left\{-\frac{1}{2}(\mathbf{y}_i - \mathbf{u}_{(j|i)l})\mathbf{\Sigma}_i^{-1}(\mathbf{y}_i - \mathbf{u}_{(j|i)l})'\right\}\right. \\ &\quad \left.\times\pi_{jl}f(Q_i = j)\right]^{\zeta_{ij}\delta_{il}}. \end{aligned}$$

Let Z_o be the generic observed data and $\mathbf{E}$ denote the generic conditional expectations conditioned on Z_o and current estimated parameters. The conditional expectation of the complete log-likelihood (omitting the constant term) is

$$\mathbf{E}\{l^f|Z_o\} = \mathbf{E}\left\{\log\prod_{i=1}^{N}\prod_{l=1}^{L}\prod_{j=0}^{2}\left[(2\pi)^{-T_i/2}|\mathbf{\Sigma}_i|^{-1/2}\right.\right.$$

$$\times \exp\left\{-\frac{1}{2}(\mathbf{y}_i-\mathbf{u}_{(j|i)l})\boldsymbol{\Sigma}_i^{-1}(\mathbf{y}_i-\mathbf{u}_{(j|i)l})'\right\}\pi_{jl}f(Q_i=j)\Big]^{\delta_{il}\zeta_{ij}}|Z_o\Bigg\}$$

$$=\sum_{i=1}^{N}\sum_{l=1}^{L}\sum_{j=0}^{2}E\delta_{il}E\zeta_{ij}\left[-\frac{T_i}{2}\log(2\pi)-\frac{1}{2}\log(|\boldsymbol{\Sigma}_i|)\right.$$

$$\left.-\frac{1}{2}(\mathbf{y}_i-\mathbf{u}_{(j|i)l})\boldsymbol{\Sigma}_i^{-1}(\mathbf{y}_i-\mathbf{u}_{(j|i)l})'+\log(\pi_{jl})+\log(f(Q_i=j))\right], \quad (2.8)$$

where $E\delta_{il}=\mathbf{E}(I(D_i=s_l)|z_{oi})$ and $E\zeta_{ij}=\mathbf{E}(I(Q_i=j)|z_{oi})$.

For the subjects whose dropout doses are observed, we have

$$E\delta_{il}=\delta_{il}, \quad (2.9)$$

$$\begin{aligned}E\zeta_{ij}&=Pr(Q_i=j|z_{oi},\boldsymbol{\Omega}_q=\boldsymbol{\Omega}_q^{[t]})\\&=\frac{\sum_{l=1}^{L}f(y_i|D_i=s_l,Q_i=j)Pr(D_i=s_l|Q_i=j)Pr(Q_i=j)I(D_i=s_l)}{\sum_{j=0}^{2}\sum_{l=1}^{L}f(y_i|D_i=s_l,Q_i=j)Pr(D_i=s_l|Q_i=j)Pr(Q_i=j)I(D_i=s_l)}\\&=\frac{\sum_{l=1}^{L}\exp\left\{-\frac{1}{2}(\mathbf{y}_i-\mathbf{u}_{(j|i)l})\boldsymbol{\Sigma}_i^{-1}(\mathbf{y}_i-\mathbf{u}_{(j|i)l})'\right\}\pi_{jl}f(Q_i=j)I(D_i=s_l)}{\sum_{j=0}^{2}\sum_{l=1}^{L}\exp\left\{-\frac{1}{2}(\mathbf{y}_i-\mathbf{u}_{(j|i)l})\boldsymbol{\Sigma}_i^{-1}(\mathbf{y}_i-\mathbf{u}_{(j|i)l})'\right\}\pi_{jl}f(Q_i=j)I(D_i=s_l)}.\end{aligned} \quad (2.10)$$

For the subjects for whom dropout doses are right censored, we have

$$\begin{aligned}E\delta_{il}&=Pr(D_i=s_l|z_{oi},\boldsymbol{\Omega}_q=\boldsymbol{\Omega}_q^{[t]})\\&=\frac{\sum_{j=0}^{2}f(y_i|D_i=s_l,Q_i=j)Pr(D_i=s_l|Q_i=j)Pr(Q_i=j)I(s_l>\tilde{D}_i)}{\sum_{l=1}^{L}\sum_{j=0}^{2}f(\mathbf{y}_i|D_i=s_l,Q_i=j)Pr(D_i=s_l|Q_i=j)Pr(Q_i=j)I(s_l>\tilde{D}_i)}\\&=\frac{\sum_{j=0}^{2}\exp\{-\frac{1}{2}(\mathbf{y}_i-\mathbf{u}_{(j|i)l})\boldsymbol{\Sigma}_i^{-1}(\mathbf{y}_i-\mathbf{u}_{(j|i)l})'\}\pi_{jl}f(Q_i=j)I(s_l>\tilde{D}_i)}{\sum_{l=1}^{L}\sum_{j=0}^{2}\exp\{-\frac{1}{2}(\mathbf{y}_i-\mathbf{u}_{(j|i)l})\boldsymbol{\Sigma}_i^{-1}(\mathbf{y}_i-\mathbf{u}_{(j|i)l})'\}\pi_{jl}f(Q_i=j)I(s_l>\tilde{D}_i)},\end{aligned} \quad (2.11)$$

$$\begin{aligned}E\zeta_{ij}&=Pr(Q_i=j|z_{oi},\boldsymbol{\Omega}_q=\boldsymbol{\Omega}_q^{[t]})\\&=\frac{\sum_{l=1}^{L}f(y_i|D_i=s_l,Q_i=j)Pr(D_i=s_l|Q_i=j)Pr(Q_i=j)I(s_l>\tilde{D}_i)}{\sum_{j=0}^{2}\sum_{l=1}^{L}f(y_i|D_i=s_l,Q_i=j)Pr(D_i=s_l|Q_i=j)Pr(Q_i=j)I(s_l>\tilde{D}_i)}\end{aligned}$$

$$= \frac{\sum_{l=1}^{L} \exp\{-(1/2)(\mathbf{y}_i - \mathbf{u}_{(j|i)l})\boldsymbol{\Sigma}_i^{-1}(\mathbf{y}_i - \mathbf{u}_{(j|i)l})'\}\pi_{jl} f(Q_i = j) I(s_l > \tilde{D}_i)}{\sum_{j=0}^{2}\sum_{l=1}^{L} \exp\{-(1/2)(\mathbf{y}_i - \mathbf{u}_{(j|i)l})\boldsymbol{\Sigma}_i^{-1}(\mathbf{y}_i - \mathbf{u}_{(j|i)l})'\}\pi_{jl} f(Q_i = j) I(s_l > \tilde{D}_i)}. \quad (2.12)$$

2.5. Observed data likelihood

The observed likelihood function, L^o, is constructed within a mixture model framework. When there is no censoring in dropout doses, we have

$$\begin{aligned} L^o(\boldsymbol{\Omega}_q) &= f(\mathbf{y}, D) \\ &= \prod_{i=1}^{N}\sum_{j=0}^{2} f(\mathbf{y}_i | D_i, Q_i = j) f(D_i | Q_i = j) f(Q_i = j) \\ &= \prod_{i=1}^{N}\prod_{l=1}^{L}\Bigg[\sum_{j=0}^{2}(2\pi)^{-T_i/2}|\boldsymbol{\Sigma}_i|^{-1/2} \\ &\quad \times \exp\Big\{-\frac{1}{2}(\mathbf{y}_i - \mathbf{u}_{(j|i)l})\boldsymbol{\Sigma}_i^{-1}(\mathbf{y}_i - \mathbf{u}_{(j|i)l})'\Big\}\pi_{jl} f(Q_i = j)\Bigg]^{\delta_{il}}. \end{aligned} \quad (2.13)$$

If there is censoring in dropout dose, we have

$$\begin{aligned} L^o(\boldsymbol{\Omega}_q) = \prod_{i=1}^{N}\Bigg[\sum_{l=1}^{L}\sum_{j=0}^{2}(2\pi)^{-T_i/2}|\boldsymbol{\Sigma}_i|^{-1/2}\exp\left\{-\frac{1}{2}(\mathbf{y}_i - \mathbf{u}_{(j|i)l})\boldsymbol{\Sigma}_i^{-1}(\mathbf{y}_i - \mathbf{u}_{(j|i)l})'\right\} \\ f(D_i = s_l | C_i, \Delta_i = 0, Q_i = j) f(Q_i = j)\Bigg], \end{aligned} \quad (2.14)$$

where

$$\begin{aligned} f(D_i = s_l | C_i, \Delta_i = 0, Q_i = j) &= f(D_i = s_l | C_i, C_i \le D_i, Q_i = j) \\ &= \frac{f(D_i = s_l | Q_i = j) I(C_i < s_l)}{\sum_{l:\{s_l > C_i\}} f(D_i = s_l | Q_i = j)}. \end{aligned} \quad (2.15)$$

2.6. Parameter estimation and variance-covariance of estimated parameters

An integrative EM and Newton-Raphson algorithm was implemented to estimate the unknown parameters $\boldsymbol{\Omega}_q$ (see the Appendix for details). After the MLEs of $\boldsymbol{\Omega}_q$ are obtained, $\text{var}(\hat{\boldsymbol{\Omega}}_q)$ is calculated. Let $S(\boldsymbol{\Omega}_q; z_o)$ and $S(\boldsymbol{\Omega}_q; z)$ denote the score vectors of observed and complete data log-likelihood functions, and $B(\boldsymbol{\Omega}_q; z_o)$ and $B(\boldsymbol{\Omega}_q; z)$ denote the negatives of the second order derivative matrices for these two types of likelihood functions, respectively. The observed

information matrix is $B(\mathbf{\Omega}_q; z_o)$, whereas the expected or Fisher information matrix is $I(\mathbf{\Omega}_q) = E_{\mathbf{\Omega}_q}(B(\mathbf{\Omega}_q; z_o))$. Asymptotically, the MLEs of parameters, $\hat{\mathbf{\Omega}}_q$, are $N(\mathbf{\Omega}_q, I^{-1}(\mathbf{\Omega}_q))$. Since $I(\hat{\mathbf{\Omega}}_q)$ and $B(\hat{\mathbf{\Omega}}_q; z_o)$ are both consistent estimators for the Fisher information matrix $I(\mathbf{\Omega}_q)$, the Fisher information matrix and, thus, the variance of the MLEs can be estimated if $I(\hat{\mathbf{\Omega}}_q)$ or $B(\hat{\mathbf{\Omega}}_q; z_o)$ is calculated.

Under regularity conditions, as described in Zacks (1971), that guarantee the log-likelihood equations can be solved and the Fisher information matrix exists, Louis (1982) proposed a method for estimating the observed information matrix of the MLEs as

$$\begin{aligned} S(\mathbf{\Omega}_q; z_o) &= \mathbf{E}(S(\mathbf{\Omega}_q; z)|z_o), \\ B(\mathbf{\Omega}_q; z_o) &= \mathbf{E}(B(\mathbf{\Omega}_q; z)|z_o) - \mathbf{E}(S(\mathbf{\Omega}_q; z)S^T(\mathbf{\Omega}_q; z)|z_o) + S(\mathbf{\Omega}_q; z_o)S^T(\mathbf{\Omega}_q; z_o). \end{aligned} \tag{2.16}$$

We have $B(\hat{\mathbf{\Omega}}_q; z_o) = \mathbf{E}(B(\hat{\mathbf{\Omega}}_q; z)|z_o) - \mathbf{E}(S(\hat{\mathbf{\Omega}}_q; z)S^T(\hat{\mathbf{\Omega}}_q; z)|z_o)$ with $\mathbf{\Omega}_q = \hat{\mathbf{\Omega}}_q$, since $S(\hat{\mathbf{\Omega}}_q; z_o) = 0$. But this method needs to calculate the conditional expectation (conditional on the observed data) of the square of the complete data scores $\mathbf{E}(S(\hat{\mathbf{\Omega}}_q; z)S^T(\hat{\mathbf{\Omega}}_q; z)|z_o)$, which is not easy. A different so-called supplemented EM algorithm or SEM algorithm was proposed by Meng and Rubin (1993) to estimate the asymptotic variance-covariance matrices; it can also be used for the calculations of the sampling errors for the MLEs of the parameters. This method avoids the calculation of the conditional expectation $E(S(\hat{\mathbf{\Omega}}_q; z)S^T(\hat{\mathbf{\Omega}}_q; z)|z_o)$, but needs to calculate the convergence rate of a "forced" EM algorithm.

Another easier method for the estimation of the variance is based on empirical Fisher information (Hogan and Laird (1997)). The sample covariance matrix of individual scores $S_i(\mathbf{\Omega}_q; z_{oi})$ is

$$\hat{I}(\mathbf{\Omega}_q; z_o) = \sum_{i=1}^{N} S_i(\mathbf{\Omega}_q; z_{oi})S_i(\mathbf{\Omega}_q; z_{oi})^T - \frac{1}{N^2}S(\mathbf{\Omega}_q; z_o)S(\mathbf{\Omega}_q; z_o)^T.$$

This is also a consistent estimator of $I(\mathbf{\Omega}_q; z_o)$. Although this method ignores the likelihood principle and can be applied to independently identically distributed cases only, it is relatively easy to compute since we need only to know individual scores. When Fisher or observed information matrices are difficult to compute, this empirical approach can be used as an alternative.

3. Hypothesis Testing

An optimal risk haplotype is selected from multiple haplotypes based on a model selection criterion. Thus, the significance of the genetic effect of the selected risk haplotype presents a multiple testing problem. Tradition approaches to correcting for multiple comparisons, such as false discovery rate control, can be used for haplotype discovery.

A major advantage of our model lies in its flexibility to generate and test a number of important hypothesis tests about the genetic control of longitudinal traits. In general, these hypotheses are sorted into the following types.

3.1. Ignorabilility test of dropout

Whether the dropout can be ignored is a first question to address, in order to better utilize the data. This can be tested by formulating the hypotheses

$$\begin{aligned} H_0 &: \boldsymbol{\Theta}_{jl} \equiv \boldsymbol{\Theta}; \pi_{jl} \equiv \pi_l, \\ H_1 &: \boldsymbol{\Theta}_{jl} \equiv \boldsymbol{\Theta}_l, \pi_{jl} \equiv \pi_l, \end{aligned} \quad \text{for } j = 2, 1, 0; l = 1, \ldots, L. \tag{3.1}$$

These hypotheses are based on a mean pattern of all subjects studied by assuming no genetic effects on longitudinal curves and dropouts. Here H_0 states that the dropout is ignorable in terms of longitudinal curves and dropout doses, whereas H_1 states that the dropout is informative, i.e., different patterns of dropout lead subjects to have different longitudinal curves.

3.2. Genetic tests with ignorable dropout

If the dropout is ignorable, by accepting the H_0 of test (3.1), we test the existence of a significant haplotype effect first on the distribution of dropout doses, and then on the distribution of longitudinal curves. The genetic effect of haplotype on the distribution of dropout doses can be tested according to

$$\begin{aligned} H_0 &: \boldsymbol{\Theta}_{jl} \equiv \boldsymbol{\Theta}; \pi_{jl} \equiv \pi_l, \\ H_1 &: \boldsymbol{\Theta}_{jl} \equiv \boldsymbol{\Theta}, \end{aligned} \quad \text{for } j = 2, 1, 0; l = 1, \ldots, L. \tag{3.2}$$

If there is a haplotype effect on the distribution of dropout doses, then whether this haplotype effect impacts longitudinal curves can be tested according to

$$\begin{aligned} H_0 &: \boldsymbol{\Theta}_{jl} \equiv \boldsymbol{\Theta}, \\ H_1 &: \boldsymbol{\Theta}_{jl} \equiv \boldsymbol{\Theta}_j, \end{aligned} \quad \text{for } j = 2, 1, 0; l = 1, \ldots, L. \tag{3.3}$$

If the null hypothesis of test (3.3) is rejected, this suggests that the haplotype has a pleiotropic effect on dropout times and longitudinal curves.

If there is no haplotype effect on the distribution of dropout doses, then whether this haplotype effect impacts longitudinal curves can be tested according to

$$\begin{aligned} H_0 &: \boldsymbol{\Theta}_{jl} \equiv \boldsymbol{\Theta}; \pi_{jl} \equiv \pi_l, \\ H_1 &: \boldsymbol{\Theta}_{jl} \equiv \boldsymbol{\Theta}_j; \pi_{jl} \equiv \pi_l, \end{aligned} \quad \text{for } j = 2, 1, 0; l = 1, \ldots, L. \tag{3.4}$$

3.3. Genetic tests with non-ignorable dropout

If the dropout is non-ignorable, by rejecting the H_0 at (3.1), we also need to test how haplotypes impact longitudinal curves and dropout doses. For the genetic effect of haplotype on the distribution of dropout doses, we test

$$\begin{aligned} &H_0 : \boldsymbol{\Theta}_{jl} \equiv \boldsymbol{\Theta}_l; \pi_{jl} \equiv \pi_l, \\ &H_1 : \boldsymbol{\Theta}_{jl} \equiv \boldsymbol{\Theta}_l, \end{aligned} \quad \text{for } j = 2, 1, 0; l = 1, \ldots, L. \tag{3.5}$$

If there is a haplotype effect on the distribution of dropout doses, then whether this haplotype effect impacts on longitudinal curves can be tested according to

$$\begin{aligned} &H_0 : \boldsymbol{\Theta}_{jl} \equiv \boldsymbol{\Theta}_l, \\ &H_1 : \text{at least one of the equalities in } H_0 \text{ does not hold,} \end{aligned} \quad \text{for } \begin{matrix} j = 2, 1, 0; \\ l = 1, \ldots, L. \end{matrix} \tag{3.6}$$

If there is no haplotype effect on the distribution of dropout doses, then whether this haplotype effect impacts longitudinal curves can be tested according to

$$\begin{aligned} &H_0 : \boldsymbol{\Theta}_{jl} \equiv \boldsymbol{\Theta}_l; \pi_{jl} \equiv \pi_l, \\ &H_1 : \boldsymbol{\Theta}_{jl} \neq \boldsymbol{\Theta}_l; \pi_{jl} \equiv \pi_l, \end{aligned} \quad \text{for } j = 2, 1, 0; l = 1, \ldots, L. \tag{3.7}$$

For all tests above, test statistics are approximately χ^2 with degrees of freedom equal to the difference in parameter number between H_1 and H_0.

4. A Worked Example

4.1. Background and data summary

The usefulness of the new model is validated by analyzing the drug response example described in the Introduction. β1AR and β2AR are candidate genes for heart function (Large et al. (1997)). In each of the two genes there are several polymorphisms common in the population. In a pharmacogenetic study comprised of 163 men and women, two SNPs at codon 49 with two alleles Ser49 (A) and Gly49 (G) and at codon 389 with two alleles Arg389 (C) and Gly389 (G) within the β1AR gene on chromosome 10, as well as two SNPs at codon 16 with two alleles Arg16 (A) and Gly16 (G) and at codon 27 with two alleles Gln27 (C) and Glu27 (G) within the β2AR gene on chromosome 5, were genotyped. A highly significant linkage disequilibrium was detected between the two SNPs for each gene ($P < 0.001$).

In this study, a drug, called dobutamine, designed to improve heart function, was injected into patients to investigate their responses in heart rate. The subjects received increasing doses of dobutamine until they achieved a target heart rate response or predetermined maximum dose. The dose levels used were

0 (baseline), 5, 10, 20, 30, and 40 mcg/min, at each of which heart rate was measured. The time interval of 3 minutes was allowed between two successive doses for subjects to reach a plateau in response to that dose. Our model is used to detect if and how haplotype variants within these candidate genes affect the response of heart rate to dobutamine. By excluding those with incomplete genetic information, we had 143 subjects involved in our analysis. Some of the subjects reached the thresholds of their heart rates before the highest dosage. About 3%, 10%, and 19% subjects dropped out at dose levels 10, 20, and 30, respectively (Figure 1). Thus, about 32% of the subjects did not complete the study.

4.2. Data analysis

4.2.1. Emax model

In a drug response experiment, there is a classical sigmoid $E_{\mathtt{max}}$ equation that can be used to describe the relationship between drug concentration (C) and drug effect (E):

$$E = E_0 + \frac{E_{\mathtt{max}} C^H}{EC_{50}^H + C^H}, \tag{4.1}$$

where E_0 is the baseline value for the drug response when the drug concentration is 0, $E_{\mathtt{max}}$ is the asymptotic (limiting) effect, EC_{50} is the drug concentration that results in 50% of the maximal effect, and H is the slope parameter that determines the slope of the concentration-response curve.

4.2.2. Data analysis by traditional functional mapping

We first analyzed the data by functional mapping, using only subjects (98) who completed the study only; this approach was used in Lin et al. (2007). Significant haplotype effects on heart rate curves were detected for two SNPs typed from gene β2AR ($p = 0.0402$), but not for two SNPs from gene β1AR. For gene β2AR, haplotype [GG] was detected to be a risk haplotype based on the likelihoods calculated by assuming that each of the four possible haplotypes, [AC], [AG], [GC], and [GG], is a risk haplotype. The next analysis was based on two SNPs from gene β2AR.

We also analyzed all subjects (143) who participated, assuming that dropout was noninformative. Thus both incompleters and completers were treated equally. This analysis did not detect any significant risk haplotype ($p = 0.0854$ for the largest likelihood under different assumptions of risk haplotype), suggesting that this treatment reduces the power of gene identification.

4.2.3. Data analysis by new functional mapping

The new model allows us to jointly model the dropout and longitudinal data. The conditional density function of the ith subject is expressed as

$$f(Y_i|D_i=s_l,Q_i=j)=(2\pi)^{-m_i/2}|\Sigma_i|^{-1/2}\exp\{-\frac{1}{2}(y_i-\mu_{ijl})\Sigma_i^{-1}(y_i-\mu_{ijl})'\}, \quad (4.2)$$

where

$$\mu_{ijl}=\left(E_{0jl}+\frac{E_{\mathtt{max}jl}t_{i1}^{H_{jl}}}{EC_{50jl}^{H_{jl}}+t_{i1}^{H_{jl}}},\ldots,E_{0jl}+\frac{E_{\mathtt{max}jl}t_{im_i}^{H_{jl}}}{EC_{50jl}^{H_{jl}}+t_{im_i}^{H_{jl}}}\right),$$

with $\{E_{0jl},E_{\mathtt{max}jl}$, and $EC_{50jl},H_{jl}\}$ being the parameters specific to subject i with composite diplotype j dropping out at s_l.

In the example of Figure 1, the dropout dose set is $S=\{s_1,s_2,s_3,s_4\}=\{10,20,30,40\}$ (Figure 1). We assumed that there was no censoring on dropout doses, $\Delta_i=1$ for $i=1,\ldots,N$. An exploratory analysis showed that for subjects who had three or four measurements, one could assume a linear or quadratic curve, respectively. For subjects who had five or six measurements, an Emax curve (4.1) was fit.

We used three schemes to jointly model dropout and longitudinal curve data: Scheme I had four differently modeled dropout patterns; Scheme II had dropouts at dose levels 10, 20, and 30 modeled in the same way, expressed as $\mathbf{\Theta}_{j1}=\mathbf{\Theta}_{j2}=\mathbf{\Theta}_{j3}$ $(j=0,1,2)$. These schemes detected haplotype [GG] as an optimal risk haplotype. Figure 2 shows the composite diplotype-specific response curves of heart rate to dobutamine under Schemes I (Figure 2A) and II (Figure 2B). The comparison of AIC values calculated between the two shows that Scheme II with 29 parameters (5980) is a better fit to the data than Scheme I with 50 parameters (6100). It is observed that composite diplotype [GG][GG] has a similar trend of heart rate for both the completers and dropouts. For this reason, we posed an additional constraint by setting composite genotype [GG][GG] equal across different dropout patterns, that is, $\mathbf{\Theta}_{j1}=\mathbf{\Theta}_{j2}=\mathbf{\Theta}_{j3}$ $(j=0,1,2)$ and $\mathbf{\Theta}_{2l}=\mathbf{\Theta}_{24}$ $(l=1,2,3)$. This model, Scheme III, is better than Scheme II in terms of AIC values. Also, since Scheme III is nested in Scheme II, we calculated the likelihood ratio of Scheme III over Scheme II, with a p-value of 0.9122 confirming the choice of Scheme III.

We will base all the subsequent analyses on Scheme III. With hypothesis test (3.1), we could test if the dropout was informative under Scheme III. The resulting likelihood ratio test statistic is $-2(-2902.4+2884.1)=36.6$, corresponding to the p-value of 2.2×10^{-7} for a χ_4^2 distribution. This suggests that dropout is informative and it is crucial to integrate this information into the analysis.

Given that the dropout is non-ignorable, we performed a hypothesis test based on test (22) to find whether risk haplotype [GG][GG] exerts a significant

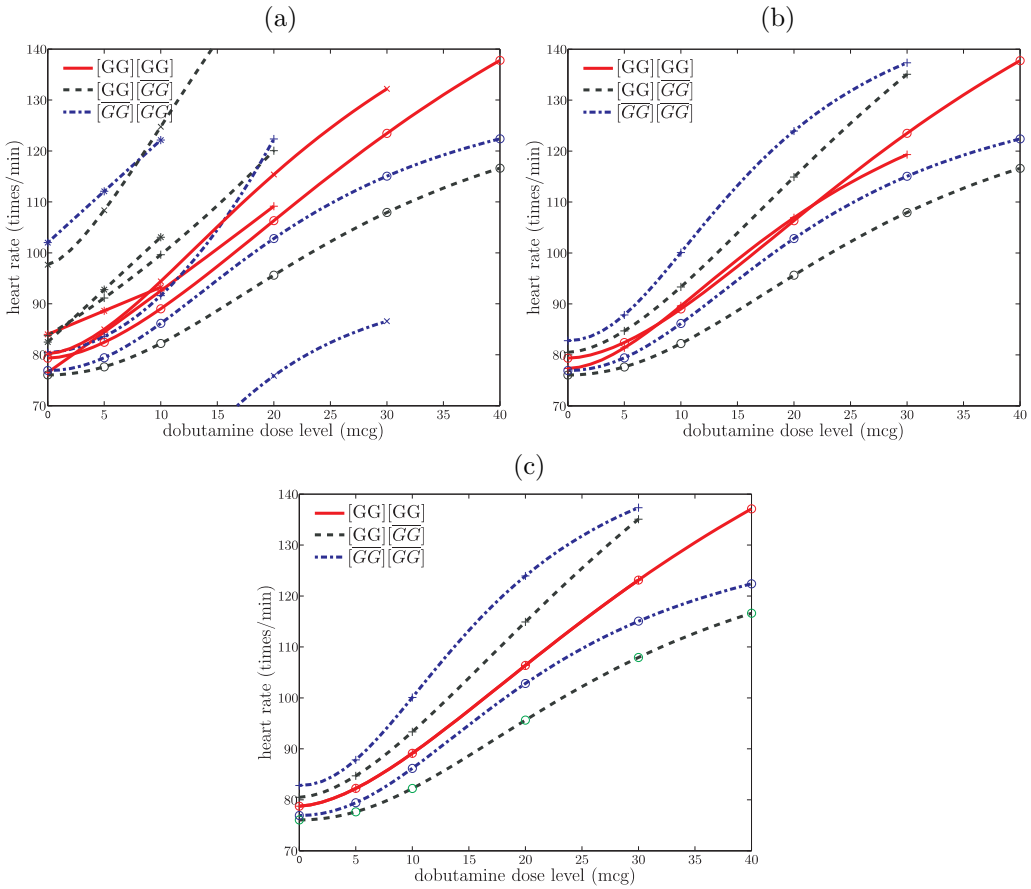

Figure 2. The response curves of heart rate to dobutamine for different composite diplotypes, [GG][GG] (solid), [GG][$\overline{\text{GG}}$] (dashed), and [$\overline{\text{GG}}$][$\overline{\text{GG}}$] (dashdot), derived from gene β2AR, under model scheme I (**A**), II (**B**), and III (**C**) (see the text).

effect on dropout doses. It was found that dropout doses do not depend on risk haplotype [GG] ($p = 0.849$). A further test for longitudinal curves based on (24) suggests that risk haplotype [GG][GG] has a significant effect on the response curve of heart rate to dobutamine ($p = 0.0329$).

It is not surprising to find that the dropouts respond to dobutamine more rapidly than the completers, but it is interesting to see that dramatic differences occur in the pattern of genetic control between the completers and dropouts (Figure 2C). In the completers, the composite diplotype with double risk haplotypes are most sensitive in drug response, followed by the composite diplotype with double non-risk haplotypes, and the composite diplotype with a risk haplotype and a non-risk haplotype; in the dropouts, the last two composite diplotypes show

Table 1. Maximum likelihood estimates (MLE) of parameters that define the model with non-ignorable dropout. Standard errors (STD) of the MLEs are estimated by Louis' approach.

Dropout Pattern	Composite Diplotype	Parameter	MLE	STD
Dropout	$Q = 0$	E_0	78.74	1.2103
		$E_{\max}$	120.33	32.0615
		EC_{50}	41.47	12.8841
		H	1.66	0.2168
	$Q = 1$	E_0	80.49	0.9548
		$E_{\max}$	131.57	57.7866
		EC_{50}	36.68	15.4237
		H	1.71	0.2545
	$Q = 2$	E_0	82.80	1.0628
		$E_{\max}$	72.58	11.3941
		EC_{50}	17.55	2.7767
		H	2.07	0.2272
Completer	$Q = 0$	E_0	78.74	1.2103
		$E_{\max}$	120.33	32.0615
		EC_{50}	41.47	12.8841
		H	1.66	0.2168
	$Q = 1$	E_0	76.18	0.6552
		$E_{\max}$	62.82	10.0181
		EC_{50}	29.53	5.0611
		H	2.00	0.2841
	$Q = 2$	E_0	76.91	0.5506
		$E_{\max}$	59.62	6.9851
		EC_{50}	22.73	2.7717
		H	2.06	0.2488
Covariance Modeling		ϕ	0.96	0.0131
		σ_e^2	62.06	1.9352
		π_1	0.3147	0.0224

a much greater sensitivity than the first one. Table 1 tabulates the MLEs of four drug response parameters $(E_0, E_{\max}, EC_{50}, H)$ for different dropout types and covariance-structuring parameters under Scheme III. As shown by the estimates of their standard errors, many parameters can be reasonably well estimated, except for $E_{\max}$.

5. Computer Simulation

5.1. Imulation scenarios

We investigated the statistical behavior of the new model using simulation studies. We simulated a population at Hardy-Weinberg equilibrium, from which a random set of samples were genotyped for a panel of SNPs and longitudinal responses at multiple dosages. We chose two associated SNPs whose haplotype and genotype frequencies (determined by allele frequencies and linkage disequilibrium) were used to simulate genotype counts in the sample. According to Liu et al. (2004), these population genetic parameters can be well estimated with the likelihood (2.1). For this reason, we did not focus on the estimates of population genetic parameters in this study.

For each subject, the phenotypic value of a longitudinal trait was simulated at nine evenly-spaced dose levels $T = (0, 5, 10, 15, 20, 25, 30, 35, 40)$, allowing a certain proportion of subjects to be dropped out non-ignorably at dose level 25. Thus, there are two possible patterns, dropouts (at dose level 25) and completers (at dose level 40), i.e., $S = s_1, s_2 = 25, 40$. The longitudinal data for each subject was simulated from a multivariate normal distribution with a Emax mean curve and a SAD(1) covariance structure. The simulation included four scenarios, depending on sample sizes, dropout rates, and heritabilities (a proportion of the phenotypic variance explained by haplotype effects):

Scenario	Sample Size	Dropout Rate (%)	Heritability
1	200	30	0.1
2	200	70	0.1
3	500	30	0.05
4	500	70	0.05

The data simulated for each scenario was analyzed using only the completers (Model 1), all the longitudinal data but ignoring dropout times (Model 2), and using the new model proposed (Model 3).

5.2. Results

For all the four scenarios, the estimates of curve parameters and SAD(1) parameters from Model 1 had the largest biases, followed by Models 2 and 3 (results not shown). Model 3 was particularly advantageous when the study had a large dropout rate. Using estimated curve parameter, we estimated the genotypic values of different composite diplotypes at all dose levels (Table 2). Because Model 1 ignores the subjects who dropout because of their high sensitivity, it tend to provide underestimates of drug response.

Also, parameter estimation from Models 2 and 3 was more precise than that from Model 1. The standard errors of the estimates of genotypic values for all

Table 2. The biases and standard errors (STD) of the estimates of genotypic values for three composite diplotypes at different dose levels under simulation scenario 1: $N = 200$, heritability of 0.1, and dropout rate of 30%

			Dose Levels								
Composite Diplotype			0	5	10	15	20	25	30	35	40
$j = 0$	Model 1	Bias	-2.05	-1.84	-3.11	-4.72	-5.94	-6.67	-7.02	-7.12	-7.06
		STD	1.15	1.42	1.57	1.67	1.73	1.70	1.65	1.69	1.89
	Model 2	Bias	0.04	-0.14	0.16	0.26	0.02	-0.42	-0.95	-1.49	-2.01
		STD	1.05	1.18	1.45	1.71	1.91	2.01	2.07	2.16	2.29
	Model 3	Bias	0.03	0.07	0.13	0.22	0.20	0.15	0.10	0.10	0.13
		STD	0.96	1.12	1.26	1.35	1.38	1.43	1.62	1.97	2.40
$j = 1$	Model 1	Bias	-0.88	-2.26	-3.82	-4.60	-4.81	-4.78	-4.68	-4.58	-4.49
		STD	0.73	0.81	1.02	1.07	1.07	1.03	1.00	1.06	1.22
	Model 2	Bias	0.02	-0.05	0.15	0.30	0.19	-0.08	-0.40	-0.71	-0.97
		STD	0.62	0.81	1.08	1.17	1.17	1.12	1.09	1.13	1.25
	Model 3	Bias	0.01	0.03	0.06	0.10	0.11	0.11	0.10	0.09	0.10
		STD	0.61	0.71	0.86	0.90	0.90	0.92	1.01	1.18	1.43
$j = 2$	Model 1	Bias	-0.73	-1.61	-2.91	-3.84	-4.36	-4.63	-4.78	-4.87	-4.92
		STD	0.97	1.04	1.31	1.41	1.44	1.46	1.42	1.43	1.60
	Model 2	Bias	0.04	-0.16	-0.18	-0.07	-0.10	-0.34	-0.72	-1.14	-1.56
		STD	0.83	0.97	1.29	1.45	1.53	1.55	1.53	1.57	1.72
	Model 3	Bias	0.03	0.04	-0.02	-0.01	0.01	0.01	0.01	0.04	0.10
		STD	0.84	0.93	1.11	1.16	1.18	1.23	1.35	1.62	2.07

different composite diplotypes from Model 3 were smaller than those from Model 2 before dose level 6 (at which some subjects dropped out). After dose level 6, Model 3 showed decreasing precision as compared to Model 2. Thus, although Model 3 is less biased, it may sacrifice some precision to reduce the bias after the dropout time.

As expected, all the models displayed increasing precision with increasing sample size and heritability, but with decreasing precision when dropout rate increased. The power for detecting significant haplotypes was also analyzed from simulation studies. Model 3 had slightly greater power than Model 2, but both had much greater power than Model 1 when the study had a large dropout rate.

6. Discussion

Integrated with genetic information collected by SNPs, functional mapping can be a useful tool to detect specific genetic variants that affect the dynamic or longitudinal patterns of outcome variables (Ma, Casella, and Wu (2002)). Func-

tional mapping has a particular power to study the pharmacogenegtic control of drug response as shown in Lin et al. (2005) and Lin et al. (2007). Functional mapping provides a quantitative framework for testing the interplay between genetic actions/interactions and the pattern of responses across different times or states. Results from functional mapping help to elucidate a comprehensive picture of a network of genetic regulations that determine the formation and progression of a disease as well as the prospective effects of drugs designed to treat the disease.

In this article, we derived a statistical model for functional mapping of longitudinal responses with non-ignorable dropout, thus broadening the implications of functional mapping to the practical setting of clinical trials. The new model was founded on the pattern mixture paradigm that considered non-ignorable missing-data mechanisms. Pattern-mixture models specify the conditional distribution of the unobserved measurements given the observed ones in a given pattern, and have been thought to be potentially useful for modeling incomplete longitudinal data (Wu and Bailey (1989); Little (1993, 1995); Fitzmaurice, Laird, and Shneyer (2001); Hogan and Laird (1997)). We incorporate the pattern-mixture model process into a framework of the mixture model for genetic mapping in which different genotypes present different curves of longitudinal responses.

The new model can not only detect the existence of specific DNA sequence variants that regulate longitudinal traits, but it also allows the tests of whether these haplotypes trigger a pleiotropic effect on longitudinal responses and the dose at which subjects drop out from the longitudinal study. The model can be further extended to consider the physiological mechanisms that cause early dropouts and model interactions between haplotypes from different gene regions and between haplotypes and environments. These modified models will find an immediate application in pharmacogenetic studies of drug response and HIV/AIDS studies in which informative dropouts commonly occur.

Acknowledgement

This work is partially supported by grants DMS/NIGMS-0540745 and NSF/IOS-0923975. We thank Prof. Julie Johnson for providing her published pharmacogenetic data to validate our new model.

Appendix

In this appendix, we describe the EM algorithm for parameter estimation. In the E step, calculate the generic conditional expectations conditional on Z_o and the current parameter estimates with equations (2.9) and (2.10), or (2.11) and

(2.12). In the M step, use these estimated expectations to solve the log-likelihood equations:

$$\frac{\partial l^o(\mathbf{\Omega}_q)}{\partial \mathbf{\Omega}_q} = \mathbf{E}\left[\frac{\partial l^f(\mathbf{\Omega}_q)}{\partial \mathbf{\Omega}_q}|Z_o\right],$$

which are specifically expressed as

$$\frac{\partial l^o(\mathbf{\Omega}_q)}{\partial \mathbf{\Theta}_j} = \mathbf{E}\left[\frac{\partial l^f(\mathbf{\Omega}_q)}{\partial \mathbf{\Theta}_j}|Z_o\right] = \sum_{i=1}^{N} E\delta_{il}E\zeta_{ij}(\mathbf{y}_i - \mathbf{u}_{(j|i)l})\mathbf{\Sigma}_i^{-1}\frac{\partial \mathbf{u}'_{(j|i)l}}{\partial \mathbf{\Theta}_j}, \tag{A.1}$$

$$\begin{aligned}\frac{\partial l^o(\mathbf{\Omega}_q)}{\partial \phi} &= \mathbf{E}\left[\frac{\partial l^f(\mathbf{\Omega}_q)}{\partial \phi}|Z_o\right] \\ &= \sum_{i=1}^{N}\sum_{l=1}^{L}\sum_{j=0}^{2} E\delta_{il}E\zeta_{ij}\Big[-\frac{1}{2}\log\frac{\partial|\Sigma_i|}{\partial\phi} \\ &\qquad -\frac{1}{2}(\mathbf{y}_i - \mathbf{u}_{(j|i)l})\frac{\partial \mathbf{\Sigma}_i^{-1}}{\partial\phi}(\mathbf{y}_i - \mathbf{u}_{(j|i)l})'\Big],\end{aligned} \tag{A.2}$$

$$\begin{aligned}\frac{\partial l^o(\mathbf{\Omega}_q)}{\partial \sigma^2} &= \mathbf{E}\left[\frac{\partial l^f(\mathbf{\Omega}_q)}{\partial \sigma^2}|Z_o\right] \\ &= \sum_{i=1}^{N}\sum_{l=1}^{L}\sum_{j=0}^{2} E\delta_{il}E\zeta_{ij}\Big[-\frac{1}{2}\log\frac{\partial|\mathbf{\Sigma}_i|}{\partial\sigma^2} \\ &\qquad -\frac{1}{2}(\mathbf{y}_i - \mathbf{u}_{(j|i)l})\frac{\partial \mathbf{\Sigma}_i^{-1}}{\partial\sigma^2}(\mathbf{y}_i - \mathbf{u}_{(j|i)l})'\Big],\end{aligned} \tag{A.3}$$

$$\begin{aligned}\frac{\partial l^o(\mathbf{\Omega}_q)}{\partial \pi_{ql}} &= \mathbf{E}\left[\frac{\partial l^f(\mathbf{\Omega}_q)}{\partial \pi_{jl}}|Z_o\right] \\ &= \sum_{i=1}^{N} E\zeta_{ij}\left[E\delta_{il}\frac{1}{\pi_{jl}} - E\delta_{iL}\frac{1}{\pi_{jL}}\right],\ l = 1,\ldots,L-1.\end{aligned} \tag{A.4}$$

To obtain the estimates of $\mathbf{\Omega}_q$ from the observed log-likelihood function $l^o(\mathbf{\Omega}_q|Z_o)$, we use the iterative steps of the Newton-Raphson algorithm:

$$\mathbf{\Omega}_q^{[t+1]} = \mathbf{\Omega}_q^{[t]} + \left\{V^{(-1)}\frac{\partial l^o}{\partial \mathbf{\Omega}_q}\right\}|_{\mathbf{\Omega}_q = \mathbf{\Omega}_q^{[t]}}, \tag{A.5}$$

where

$$V = -\frac{\partial^2 l^o(\mathbf{\Omega}_q)}{\partial \mathbf{\Omega}_q \partial \mathbf{\Omega}_q'}$$

is the negative of the second order derivatives matrix, or the observed information matrix. In practical Newton Raphson iterations, V is replaced by

$$V_1 = \mathbf{E}\left[-\frac{\partial^2 l^f(\mathbf{\Omega}_q)}{\partial \mathbf{\Omega}_q \partial \mathbf{\Omega}_q'}|Z_o\right],$$

which is always larger than V.

The general Newton-Raphson algorithm converges quickly if appropriate initial values are selected for the parameters. An efficient procedure uses the results from the simplex algorithm as initial values.

References

Akey, J., Jin, L. and Xiong, M. (2001). Haplotypes vs single marker linkage disequilibrium tests: what do we gain? *Eur. J. Hum. Genet.* **9**, 291-300.

Bader, J. S. (2001). The relative power of snps and haplotype as genetic markers for ssociation tests. *Pharmacogenomics* **2**, 11-24.

Collins, F. S., Guyer, M. S. and Charkravarti, A. (1997). Variations on a theme: cataloging human DNA sequence variation. *Science* **278**, 1580-1581.

Fitzmaurice, G. M., Laird, N. M. and Shneyer, L. (2001). An alternative parameterization of the general linear mixture model for longitudinal data with non-ignorable drop-outs. *Statist. Medicine* **20**, 1009-1021.

Hogan, J. W. and Laird, N. M. (1997). Model-based approaches to analysing incomplete longitudinal and failure time data. *Statist. Medicine* **16**, 259-272.

Huang, B. E., Amos, C. I. and Lin, D. Y. (2007). Detecting haplotype effects in genomewide association studies. *Genet. Epidemiol.* **31**, 803-812.

Judson, R., Stephens, J. C. and Windemuth, A. (2000). The predictive power of haplotypes in clinical response. *Pharmacogenomics* **1**, 15-26.

Large, V., Hellstrom, L., Reynisdottir, S., Lonnqvist, F., Eriksson, P., Lannfelt, P. and Arner P. (1997). Human beta-2 adrenoceptor gene polymorphisms are highly frequent in obesity and associate with altered adipocyte beta-2 adrenoceptor function. *J. Clinical Investigations* **100**, 3005-3013.

Lin, D. Y. and Huang, B. E. (2007). The use of inferred haplotypes in downstream analyses. *Am. J. Hum. Genet.* **80**, 577-579.

Lin, D. Y. and Zeng, D. (2006). Likelihood-based inference on haplotype effects in association studies (with discussion). *J. Amer. Statist. Assoc.* **101**, 89-118.

Lin, M., Aquilante, C., Johnson, J. A. and Wu, R. L. (2005). Sequencing drug response with HapMap. *The Pharmacogenomics J.* **5**, 149-156.

Lin, M., Hou, W., Li, H. Y., Johnson, J. A. and Wu, R. L. (2007). Modeling interactive quantitative trait nucleotides for drug response. *Bioinformatics* **23**, 1251-1257.

Little, R. J. A. (1993). Pattern-mixture models for multivariate incomplete data. *J. Amer. Statist. Assoc.* **88**, 125-134.

Little, R. J. A. (1995). Modeling the dropout mechanism in repeated-measures studies. *J. Amer. Statist. Assoc.* **90**, 1112-1121.

Liu, T., Johnson, J. A., Casella, G. and Wu, R. L. (2004). Sequencing Complex Diseases With HapMap. *Genetics* **168**, 503-511.

Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *J. Roy. Statist. Soc. Ser. B (Methodological)* **44**, 226-233.

Lynch, M. and Walsh, J. B. (1998). *Genetics and Analysis of Quantitative Traits*, Sinauer Assocs., Inc., Sunderland, MA.

Ma, C. X., Casella, G. and Wu, R. L. (2002). Functional mapping of quantitative trait loci underlying the character process: a theoretical framework. *Genetics* **161**, 1751-1762.

Meng, X. L. and Rubin, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: a general framework . *Biometrika* **80**, 267-278.

Morris, R. W. and Kaplan, N. L. (2002). On the advantage of haplotype analysis in the presence of multiple disease susceptibility alleles. *Genetic Epidemiology* **23**, 221-233.

Rha, S. Y., Jeung, H. C., Choi, Y. H., Yang, W. I., Yoo, J. H., Kim, B. S., Roh, J. K. and Chung, H. C. (2007). An association between RRM1 haplotype and gemcitabineinduced neutropenia in breast cancer patients. *Oncologist* **12**, 622-630.

Wu, M. C. and Bailey, K. R. (1989). Estimation and comparison of changes in the presence of informative right censoring: conditional linear model. *Biometrics* **45**, 939-955.

Wu, R. L. and Lin, M. (2006). Functional mapping : A new tool to study the genetic architecture of dynamic complex traits. *Nature Reviews Genetics* **7**, 229-237.

Zimmerman, D. L. and Núñez-Antón, V. (2001). Parametric modeling of growth curve data: An overview (with discussion). *Test* **10**, 1-73.

Zacks, S. (1971). *The Theory of Statistical Inference*, Wiley, New York.

Zaykin, D. V., Westfall, P. H., Young, S. S., Karnoub, M. A., Wagner, M. J. and Ehm, M. G. (2002). Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. *Human Heredity* **53**, 79-91.

Department of Statistics, University of Florida, Gainesville, FL 32611, U.S.A.

E-mail: hongying.li.joy@gmail.com

Center for Statistical Genetics, Pennsylvania State University, Hershey, PA 17033, U.S.A.

E-mail: rwu@hes.hmc.psu.edu

(Received November 2009; accepted November 2010)