AN EFFICIENT AND COMPUTATIONALLY ROBUST STATISTICAL METHOD FOR ANALYZING CASE-CONTROL MOTHER–OFFSPRING PAIR GENETIC ASSOCIATION STUDIES

BY HONG ZHANG[1], BHRAMAR MUKHERJEE[2], VICTORIA ARTHUR[3,*], GANG HU[3,**], HAGIT HOCHNER[4] AND JINBO CHEN[3,†]

[1]*Department of Statistics and Finance, School of Management, University of Science and Technology of China, zhangh@ustc.edu.cn*

[2]*Department of Biostatistics, University of Michigan, bhramar@umich.edu*

[3]*Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, *vlynn@pennmedicine.upenn.edu; **Gang.Hu@pbrc.edu; †jinboche@mail.med.upenn.edu*

[4]*The Hebrew University of Jerusalem, hagith@ekmd.huji.ac.il*

Case-control mother–offspring pair design has been widely adopted for studying early-life and women's pregnancy health. It allows assessment of pre- and perinatal environmental risk factors as well as both maternal and offspring genetic risk factors. Data arising from this design is routinely analyzed using standard prospective logistic regression. Such data has two unique features: the offspring genotypes are not correlated with maternal environmental risk factors given maternal genotypes, and offspring and maternal genotypes are related through mendelian transmission. In this work, built upon a novel regression model relating maternal genotypes to environmental risk factors, we proposed a novel retrospective likelihood method that effectively utilized the two data features to increase statistical efficiency for detecting maternal and offspring genetic effects. The inference procedure was based on a profile likelihood derived using the Lagrange multiplier method, but we replaced the multipliers with their large sample limits to enable highly efficient and computationally stable estimation. We showed that our proposed estimates of odds ratio association parameters are consistent and asymptotically normally distributed and demonstrated the finite sample performance through extensive simulation studies and application to genetic association studies of birth weight and gestational diabetes mellitus.

1. Introduction. Emerging evidence suggests that obstetrical and early life phenotypes can be altered by environmental factors and genes of both the mother and offspring (Kanayama et al. (2002), Saftlas, Beydoun and Triche (2005), Wangler et al. (2005), Goddard et al. (2007)). For example, the Jerusalem Perinatal Study (JPS; Harlap et al. (2007)) recently reported that maternal prepregnancy BMI and both maternal and offspring genes may be implicated in the risk of low birth weight. Because offspring genotypes may affect maternal physiology during pregnancy (Petry, Ong and Dunger (2007)), case-control studies of maternal phenotypes, such as premature birth, preeclampsia and gestational diabetes mellitus (GDM), have started to routinely collect risk factor information from both mothers and offspring. Cases and controls can be mothers, if the phenotype is maternal, or offspring, if the phenotype is for an early-life condition. Standard prospective logistic regression analysis (Prentice and Pyke (1979)) can be applied to analyze such case-control mother–offspring pair data. Alternative statistical methods have recently become available that have increased efficiency for estimating odds ratio (OR) association parameters. In a loglinear model framework, efficiency improvement was achieved through incorporation of family information,

Key words and phrases. Case-control mother–offspring pair design, genetic association, profile likelihood, saddle point problem, retrospective likelihood.

including mendelian inheritance between maternal and offspring genotypes, genetic mating symmetry and parental allelic exchangeability (Shi et al. (2008)). A retrospective likelihood method was developed to fit logistic regression models of maternal-offspring genetic effects with incorporation of mendelian inheritance and Hardy–Weinberg equilibrium (HWE) (Chen, Zheng and Wilson (2009)) and later extended to allow assessment of environmental effects and gene-environment interactions through a semiparametric maximum likelihood estimation (MLE) method (Chen, Lin and Hochner (2012)). Improvement in efficiency for the semiparametric MLE method was also due to an additional constraint reflecting that environmental factors from the maternal source are naturally conditionally independent of the offspring genotype given the maternal genotype. The power for testing genetic and gene-environment interaction effects can be greatly improved by this method. This semiparametric MLE method requires enumeration of case and control mother–offspring pairs in the cohort from which subjects were sampled, and it can have numerical difficulty because maximization of the likelihood involves escape of saddle points, which is often unsuccessful.

In the current work we propose a novel restrospective likelihood method for analyzing case-control mother–offspring pair data. Our method is statistically efficient, computationally stable and widely applicable to any case-control studies without requiring that cases and controls be selected from a cohort that investigators can enumerate. The rest of the paper is organized as follows. In Section 2 we describe a new retrospective likelihood function for case-control mother–offspring pair data under a logistic penetrance model. A novel estimation and testing method is developed using the profile likelihood method. We illustrate the proposed method through simulation studies in Section 3 and apply it to two real datasets in Section 4. We make some concluding remarks in Section 5.

2. Methods.

2.1. *Notation, model and prospective likelihood.* Let Y denote case-control status ($Y = 1$: case; $Y = 0$: control), X a vector of p risk factors (can include some covariates), G^o the offspring genotype of an autosomal single nucleotide polymorphism (SNP) and G^m the maternal genotype of the same SNP. Let A and a denote the two alleles of the SNP with a denoting the minor allele with frequency ≤ 0.5. Then, the genotype can be coded using two dummy variables. The first takes value 1 if the genotype is Aa and 0 otherwise, and the second one takes 1 if the genotype is aa and 0 otherwise. When a mode of inheritance is assumed, a genotype can be coded using one single variable. For the additive mode of inheritance, a genotype is coded as the minor allele count divided by 2 (0 for AA, 0.5 for Aa, and aa for 1). For the dominant mode of inheritance, a genotype is coded as 1 if at least one copy of a is present and 0 otherwise. For the recessive mode of inheritance, a genotype is coded as 1 if it is aa and 0 otherwise. The data available for analysis, denoted as (Y_i, G_i^o, G_i^m, X_i), is collected from n_1 cases and n_0 controls, where subscript i indexes the ith subject ($i = 1, \ldots, n_1 + n_0$). Let n denote the total number of subjects, that is, $n = n_1 + n_0$. The relationship between Y and (G^o, G^m, X) is described by a logistic penetrance model

$$\text{pr}(Y = 1|G^o, G^m, X) = \text{expit}\{\boldsymbol{\beta}^T g(G^o, G^m, X)\}, \tag{2.1}$$

where T denotes the vector transpose, expit$(\cdot)$ is defined as $e^{\cdot}/(1 + e^{\cdot})$, $g(G^o, G^m, X)$ is a user-specified function of risk factors (G^o, G^m, X) and $\boldsymbol{\beta}$ denotes the vector of regression parameters. Because the joint effect of risk factors (G^o, G^m, X) can be complex, we use a general functional form $g(\cdot)$. It may be of interest to consider only maternal or offspring genetic effect. Maternal and offspring genes may interact with each other, and both maternal and offspring genes may modify effects of environmental risk factors. When no mode of inheritance is assumed, the most general penetrance model can involve four main genetic

effects, four gene-gene interaction effects, p main covariate effects and $4p$ gene-environment interactions. Specification of a mode of inheritance effectively reduces the dimension of regression parameters. For the dominant/additive/recessive mode of inheritance, some model examples of $\boldsymbol{\beta}^\tau g(G^o, G^m, X; \boldsymbol{\beta})$ are given as follows:

$$
\begin{aligned}
&\text{M1:} \quad \boldsymbol{\beta}^\tau g(G^o, G^m, X) = \beta_0 + \beta_{G^o} G^o + \beta_X^\tau X,\\
&\text{M2:} \quad \boldsymbol{\beta}^\tau g(G^o, G^m, X) = \beta_0 + \beta_{G^m} G^m + \beta_X^\tau X,\\
&\text{M3:} \quad \boldsymbol{\beta}^\tau g(G^o, G^m, X) = \beta_0 + \beta_{G^o} G^o + \beta_X^\tau X + \beta_{G^o X}^\tau G^o X,\\
&\text{M4:} \quad \boldsymbol{\beta}^\tau g(G^o, G^m, X) = \beta_0 + \beta_{G^m} G^m + \beta_X^\tau X + \beta_{G^m X}^\tau G^m X,\\
&\text{M5:} \quad \boldsymbol{\beta}^\tau g(G^o, G^m, X) = \beta_0 + \beta_{G^o} G^o + \beta_{G^m} G^m + \beta_X^\tau X\\
&\qquad\qquad\qquad\qquad + \beta_{G^o X}^\tau G^o X + \beta_{G^m X}^\tau G^m X.
\end{aligned}
\tag{2.2}
$$

Note that X, $G^o X$ and $G^o X$ are vectors of the same length since both G^o and G^m are scalars with the mode of inheritance incorporated. Models M1 and M2 involve only main effects of environmental risk factors and the genotype of only the mother or offspring. Models M3 and M4 have additional terms for interaction effects between environmental risk factors and the genotype. Model M5 allows full assessment of the joint effect of maternal and offspring genotypes and their respective interaction with environmental risk factors. Denote by $\boldsymbol{\beta}_G$ the parameter vector consisting of all genetic related effects, for instance, $\boldsymbol{\beta}_G = (\beta_{G^o}, \beta_{G^m}, \beta_{G^o X}^\tau, \beta_{G^m X}^\tau)^\tau$ in model M5. This work considers efficient estimation of $\boldsymbol{\beta}_G$ and powerful testing of hypotheses on $\boldsymbol{\beta}_G$. For example, in model M5 one may test a global null hypothesis of absence of any genetic effect, $\boldsymbol{\beta}_G = 0$, or absence of any interaction effect, $(\beta_{G^o X}^\tau, \beta_{G^m X}^\tau)^\tau = 0$ or absence of the maternal genetic effect, $(\beta_{G^m}, \beta_{G^m X}) = 0$.

2.2. *A novel restrospective likelihood function*. We base our inference on maximizing the retrospective likelihood function

$$\prod_{i=1}^n \operatorname{pr}(G_i^o, G_i^m, X_i | Y_i),$$

under the constraints of external information on phenotype prevalence, $\operatorname{pr}(Y = 1) = f$, mendelian relationship between maternal and offspring genotypes G^o and G^m and conditional independence between G^o and X given G^m, $\operatorname{pr}(X|G^o, G^m) = \operatorname{pr}(X|G^m)$. The second and third constraints, which reflect unique features of case-control mother–offspring pair data, were also adopted in Chen, Lin and Hochner (2012). By Bayes' theorem the inference problem is equivalent to maximizing the likelihood function $\prod_{i=1}^n \operatorname{pr}(Y_i, G_i^o, G_i^m, X_i)/f$ under the same constraints. Note that f is specified a priori, so that this likelihood function is equivalent to

$$\prod_{i=1}^n \operatorname{pr}(Y_i, G_i^o, G_i^m, X_i)$$

under the same constraints.

Let π_i denote $\operatorname{pr}(X_i)$, the probability mass of X placed at the observed value X_i and let $\boldsymbol{\pi}$ denote the vector of π_i's, $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_n)$. The joint probability $\operatorname{pr}(G_i^o, G_i^m, X_i)$ can be written as

$$\operatorname{pr}(G_i^o, G_i^m, X_i) = \operatorname{pr}(G_i^o | G_i^m) \operatorname{pr}(G_i^m | X_i) \pi_i$$

under the conditional independence between G^o and X given G^m. When the mendelian inheritance law and random mating hold in the parental population, the conditional distribution

TABLE 1

The distribution of the offspring genotype conditional on the maternal genotype under various modes of inheritance (additive, dominant and recessive). Here, θ is the MAF, and $D = F/(1-F)$ with F being the fixation index parameter

Additive

	G^o		
G^m	0	$\frac{1}{2}$	1
0	$1-\theta$	θ	0
$\frac{1}{2}$	$\frac{1-\theta}{2}$	$\frac{1}{2}$	$\frac{\theta}{2}$
1	0	$1-\theta$	θ

Dominant

	G^o	
G^m	0	1
0	$1-\theta$	θ
1	$\frac{(1-\theta)^2}{2-\theta+D}$	$1-\frac{(1-\theta)^2}{2-\theta+D}$

Recessive

	G^o	
G^m	0	1
0	$1-\frac{\theta^2}{1+\theta+D}$	$\frac{\theta^2}{1+\theta+D}$
1	$1-\theta$	θ

$\text{pr}(G_i^o|G_i^m)$ is a function of the MAF θ and fixation index parameter F (Table 1). Here, F is a measure of deviation from HWE in the maternal population, and F equals zero if and only if HWE holds. To fully specify our likelihood function, we propose a novel parametric regression model to quantify the relationship between G^m and X, where G^m adopts the same numerical coding as above (2.1):

$$\text{pr}(G^m = k|X) = \frac{\xi_k(\theta, F)\exp\{k\eta^\tau X)\}}{\sum_l \xi_l(\theta)\exp\{l\eta^\tau X)\}}. \tag{2.3}$$

We will specify function $\xi_k(\theta, F)$ shortly, and η is a vector of regression parameters. We call model (2.3) a "double-additive" logistic regression ("daLOG") model because the logarithm of OR function is additive in both k and X. That is, for an increment Δ_X in X and an increment Δ_k in k, the OR function

$$\frac{\text{pr}(G^m = k+\Delta_k|X+\Delta_X)\,\text{pr}(G^m = k|X)}{\text{pr}(G^m = k|X+\Delta_X)\,\text{pr}(G^m = k+\Delta_k|X)}$$

is equal to $\exp\{\Delta_k\eta^\tau\Delta_X\}$ which is free of both X and k. The function $\xi_k(\theta, F)$ is determined by the MAF θ and fixation index parameter F, and we give the forms of $\xi_k(\theta, F)$ in Table 2. For example, under the recessive mode of inheritance, $\xi_1(\theta, F) = \text{pr}(G_i^m = 1) = (1-F)\theta^2 + F\theta$ and $\xi_0(\theta, F) = \text{pr}(G_i^m = 0 \text{ or } G_i^m = 0.5) = 1 - \text{pr}(G_i^m = 1) = 1 - (1-F)\theta^2 - F\theta$. Note that the model (2.3) has no intercept, and $\text{pr}(G^m = k|X = 0) = \xi_k(\theta, F) = \text{pr}(G^m = k)$. That is, $\text{pr}(G^m = k|X = 0)$ reduces to the marginal probability $\text{pr}(G^m = k)$.

Let Θ denote the vector of all unknown model parameters $(\theta, F, \boldsymbol{\beta}^\tau, \eta^\tau)^\tau$. The empirical likelihood function under the above-mentioned constraints can be written as

$$L(\Theta, \boldsymbol{\pi}) = \prod_{i=1}^{n} \text{pr}(Y_i|G_i^o, G_i^m, X_i)\,\text{pr}(G_i^o|G_i^m)\,\text{pr}(G_i^m|X_i)\pi_i. \tag{2.4}$$

In the next subsection we derive the profile likelihood function for Θ and discuss the challenge in maximizing it to obtain an estimate of Θ. Then, we describe a modification to the

TABLE 2

Specification of $\xi_k(\theta, F)$ in the daLOG model for various mode of inheritance. Here, θ is the MAF, $\tilde{\theta} = 1 - \theta$, and F is the fixation index parameter

	$k = 0$	$k = 1/2$	$k = 1$
Recessive	$1 - (1 - F)\theta^2 - F\theta$	–	$(1 - F)\theta^2 + F\theta$
Additive	$(1 - F)\tilde{\theta}^2 + F(1 - \theta)$	$2(1 - F)\theta\tilde{\theta}$	$(1 - F)\theta^2 + F\theta$
Dominant	$(1 - F)\tilde{\theta}^2 + F\tilde{\theta}$	–	$1 - (1 - F)\tilde{\theta}^2 - F\tilde{\theta}$

profile likelihood function that leads to a computationally simple but efficient estimation procedure.

It is noted that the daLOG model is applicable to any genetic mode of inheritance (recessive, additive or dominant) specified at the beginning of Section 2.1. The daLOG model bears resemblance to the baseline logit (Agresti (2013)) and proportional odds model for a polytomous outcome, but it is more parsimonious in specifying the effect of X with the "intercept" parameter(s) being a function of the MAF θ. The daLOG model involves one single log-OR parameter vector η and the MAF θ as a nuisance parameter which can be estimated by utilizing information from both $\text{pr}(G_i^m|X_i)$ and $\text{pr}(G_i^o|G_i^m)$. Under the additive mode of inheritance where k takes values 0, 0.5 and 1, the baseline logit model would involve two intercept parameters and two log-OR parameter vectors for the X effect, and the proportional odds model would involve two intercept parameters and one log-OR parameter vector for the X effect. Therefore, statistical inference on the parameter vector β_G is expected to be less efficient under these two existing models. Under the dominant or recessive mode of inheritance, the daLOG model reduces to a logistic regression model form but with the intercept being a function of θ. Therefore, our daLOG model can be seen as a generalization of the conventional logistic regression model for relating G^m and X. As we make clear later, factorization of the nuisance distribution $\text{pr}(X, G^m)$ as $\text{pr}(G^m|X)\,\text{pr}(X)$ and application of the daLOG model (2.3) to $\text{pr}(G^m|X)$ are the key innovation for us to modify the profile likelihood function to avoid similar computation difficulty as experienced by the semiparametric MLE of Chen, Lin and Hochner (2012). Note the latter factorized $\text{pr}(X, G^m)$ as $\text{pr}(X|G^m)\,\text{pr}(G^m)$ and kept $\text{pr}(X|G^m)$ as nonparametric.

2.3. *A semiparametric method based on a modified profile likelihood.* Using the Lagrange multiplier method, we first profile $\boldsymbol{\pi}$ out of the likelihood function $L(\Theta, \boldsymbol{\pi})$ given in (2.4) under the constraint $\text{pr}(Y = 1) = f$. As derived in Appendix A (Zhang et al. (2020)), the log profile likelihood function of Θ is equal to

$$l_{\text{p}}(\Theta) \equiv l(\Theta, \lambda_\Theta), \tag{2.5}$$

where $l(\Theta, \lambda)$ is defined as

$$\begin{aligned} l(\Theta, \lambda) = &\sum_{i=1}^{n} \log \text{pr}\big(Y_i|G_i^o, G_i^m, X_i\big) + \sum_{i=1}^{n} \log \text{pr}\big(G_i^o|G_i^m\big) + \sum_{i=1}^{n} \log \text{pr}\big(G_i^m|X_i\big) \\ &- \sum_{i=1}^{n} \log\big[n\big\{1 + \lambda\big(H_i(\Theta) - f\big)\big\}\big], \end{aligned}$$

and λ_Θ is the solution to the following equation with respect to λ,

$$\sum_{i=1}^{n} \frac{H_i(\Theta) - f}{1 + \lambda(H_i(\Theta) - f)} = 0. \tag{2.6}$$

Here,

$$H_i(\Theta) = \sum_j \sum_k \{\text{pr}(Y = 1|G^o = j, G^m = k, X_i) \times \text{pr}(G^o = j|G^m = k)\,\text{pr}(G^m = k|X_i)\}.$$

Note that equation (2.6) is equivalent to the "score" equation

$$\frac{\partial l(\Theta, \lambda)}{\partial \lambda} = 0.$$

Thus, the maximum profile likelihood estimator of Θ can be obtained by jointly solving the "score" equations

$$\frac{\partial l(\Theta, \lambda)}{\partial \lambda} = 0 \quad \text{and} \quad \frac{\partial l(\Theta, \lambda)}{\partial \Theta} = 0. \tag{2.7}$$

However, $l(\Theta, \lambda)$ is not a true log likelihood function since λ is a constructed parameter. In the special scenario of standard case-control genetic association studies where no genetic data from family members is available, the current profile likelihood function (2.5) coincides with that derived in Zhang et al. (2018) which was shown to be convex with respect to λ. In general, it is computationally unstable to solve the "score" equations since the solution is a saddle point of $l(\Theta, \lambda)$. To resolve the numerical problem, we propose to modify the profile likelihood function (2.5) by replacing λ_Θ with its limiting value, λ_0, following an idea developed in Zhang et al. (2018). Denote by Θ_0 the true value of Θ. The "true" value λ_0 is defined as the solution to the equations

$$E\left[\frac{\partial l(\Theta, \lambda)}{\partial \Theta}\right]\Bigg|_{\Theta=\Theta_0, \lambda=\lambda_0} = 0, \qquad E\left[\frac{\partial l(\Theta, \lambda)}{\partial \lambda}\right]\Bigg|_{\Theta=\Theta_0, \lambda=\lambda_0} = 0. \tag{2.8}$$

It turns out that λ_0 has a closed form,

$$\lambda_0 = \frac{n_1}{nf} - \frac{n_0}{n(1-f)}. \tag{2.9}$$

The readers are referred to Appendix B (Zhang et al. (2020)) for the detailed proof. These results were obtained by extending the proof of Theorem 1 in Zhang et al. (2018), where the maternal genotype G^m was not involved and G^o and X were assumed to be independent. Specifically, the current proof needed to account for the correlation between maternal and offspring genotypes and between G^m and X.

The two equations in (2.8) are essential for deriving a computationally simple and stable inference method involving λ_0, as stated below. Note that the second equation in (2.8) corresponds to (2.6) which holds under the likelihood function (2.4). On the other hand, no analogue to (2.6) exists that can make (2.8) hold when the nuisance distribution $\text{pr}(X, G^m)$ is factorized into $\text{pr}(X|G^m)\,\text{pr}(G^m)$ as in Chen, Lin and Hochner (2012). This theoretically motivated the daLOG model and the corresponding likelihood function (2.4). We define the modified profile likelihood function as

$$l_{\text{mp}}(\Theta) := l(\Theta, \lambda_0), \tag{2.10}$$

where $\lambda(\Theta)$ in (2.5) was replaced by λ_0, the limit value of $\lambda(\Theta)$. We propose a new estimator of Θ, denoted by $\hat{\Theta}_{\text{p}}$, by solving the corresponding score equation

$$\frac{\partial l_{\text{mp}}(\Theta)}{\partial \Theta} = 0. \tag{2.11}$$

Because Θ is of finite dimension, $\hat{\Theta}_{\mathrm{p}}$ can be obtained by simply using the Newton–Raphson algorithm. Its asymptotic properties can be established using the standard large sample theory for Z-statistic. Define matrices $A(\Theta_0)$ and $\Sigma(\Theta_0)$ as

$$A(\Theta_0) = \frac{1}{n} E\left[\frac{\partial^2 l_{\mathrm{mp}}(\Theta)}{\partial\Theta\partial\Theta^{\tau}}\right]\Bigg|_{\Theta=\Theta_0} \quad \text{and} \quad \Sigma(\Theta_0) = \frac{1}{n}\operatorname{cov}\left(\frac{\partial l_{\mathrm{mp}}(\Theta)}{\partial\Theta}\right)\Bigg|_{\Theta=\Theta_0}.$$

We can show that with a probability tending to one as $n \to \infty$, a solution $\hat{\Theta}_{\mathrm{p}}$ to equation (2.11) exists, and $\hat{\Theta}_{\mathrm{p}}$ is a consistent estimator of Θ (i.e., $\hat{\Theta}_{\mathrm{p}} \xrightarrow{P} \Theta$). In addition, $\hat{\Theta}_{\mathrm{p}}$ is asymptotically normally distributed,

$$\sqrt{n}(\hat{\Theta}_{\mathrm{p}} - \Theta) \xrightarrow{D} N\{0, A^{-1}(\Theta_0)\Sigma(\Theta_0)A^{-1}(\Theta_0)\}.$$

Obviously, $A(\Theta_0)$ can be consistently estimated by the empirical counterpart

$$\hat{A} := \frac{1}{n}\frac{\partial^2 l_{\mathrm{mp}}(\Theta)}{\partial\Theta\partial\Theta^{\tau}}\Bigg|_{\Theta=\hat{\Theta}}.$$

Note that $\partial l_{\mathrm{mp}}(\Theta)/\partial\Theta$ is the summation of two sets of independent random vectors separately for the case and control mother–offspring pairs, and the random vectors in each set are identically distributed. Therefore, $\Sigma(\Theta_0)$ can be consistently estimated by $\hat{\Sigma}$, the summation of two corresponding sample variance-covariance matrices multiplied by the respective numbers of cases and controls divided by the total sample size n. Now, the limiting variance of $\sqrt{n}\hat{\Theta}_{\mathrm{p}}$ can be consistently estimated by $\hat{A}^{-1}(\hat{\Theta})\hat{\Sigma}(\hat{\Theta})\hat{A}^{-1}(\hat{\Theta})$, and Wald tests can be consequently constructed for testing the genetic effects. Furthermore, following Zhang et al. (2018), we can show that $\hat{\Theta}$ has the same efficiency as the MLE that maximizes the profile likelihood function (2.5) (Zhang et al. (2020), Appendix C).

Our method exploited a natural fact in studies of obstetrical and early-life outcomes where offspring genotypes are uncorrelated with maternal environmental risk factors given maternal genotypes. In line with the literature on exploiting gene-environment independence to increase statistical efficiency for assessing gene-environment interaction effects (Piegorsch, Weinberg and Taylor (1994), Umbach and Weinberg (1997), Chatterjee and Carroll (2005)), a special case of our method is to enforce independence also between maternal genotypes and environmental risk factors. This can be achieved by simply setting $\eta = 0$ in the daLOG model. While this additional constraint may lead to further efficiency improvement, the corresponding estimator is biased when this additional constraint is not satisfied (Chatterjee and Carroll (2005)). We will assess the bias and variance tradeoff of this estimator below in simulation studies.

2.4. *Implementation of the proposed method.* We have developed an R package called "CCMO," which is abbreviation for "case-control mother–offspring," to implement our new method. CCMO has been made available at GitHub (http://github.com/zhanghfd/CCMO), a free web-based Git repository hosting service. CCMO provides three main functions. The first function, `singleSNP`, estimates and/or tests unknown parameters (log-odds ratios β and η, MAF θ and fixation index parameter F) for a single SNP; The second function, `OmnibusTest`, simultaneously tests multiple effects with a Wald statistic; The third function, `multipleSNP`, analyzes multiple SNPs by allowing utilization of multiple CPU cores to reduce computation time. The input of CCMO includes genotypes of mother–offspring pairs, case-control status of mothers or offspring and covariates of the mother. Users can flexibly specify the design matrices associated with the penetrance function and the daLOG model. Three modes of inheritance (additive, dominant and recessive) can be incorporated. HWE ($F = 0$) and/or independence between genotypes and covariates ($\eta = 0$) can be enforced in this R package.

3. Simulation studies. For convenience, we henceforth refer to our modified profile likelihood method under model (2.3) as "DEP," where "DEP" is abbreviation for "dependence" to emphasize that our method allows dependence between maternal genotype G^m and environmental risk factors X. Similarly, we refer to the semiparametric maximum likelihood method proposed by Chen, Lin and Hochner (2012) as "DEP-CLH" and to an analogue of DEP with an additional constraint of independence between G^m and X (i.e., set $\eta = 0$ in the daLOG model) as "IND" (abbreviation for "independence"). DEP and DEP-CLH are not strictly comparable, because one requires known phenotype prevalence and the other requires enumeration of the underlying sampling cohort. But some comparison would inform pros and cons of our method related to specification of the daLOG model and modification of the profile likelihood. To this end, we here consider DEP-CLH as a slightly modified version of the Chen, Lin and Hochner (2012) method as follows. Consider case-control mother–offspring pair data collected from an enumerated cohort for which the case-control status is known for all subjects in a prospective cohort of size N ($N > n$). The likelihood function in Chen, Lin and Hochner (2012) was written as

$$\left\{\prod_{i=1}^{n} \operatorname{pr}(Y_i|G_i^o, G_i^m, X_i)\operatorname{pr}(X_i|G_i^m)\operatorname{pr}\big(G_i^o, G_i^m\big)\right\} \times \{\operatorname{pr}(Y=0)^{M_0}\operatorname{pr}(Y=1)^{M_1}\}, \tag{3.1}$$

where M_1 and M_0 denote the respective number of cases and controls in the cohort who were not included in the case-control sample ($M_1 + M_0 = N - n$). Note that $\operatorname{pr}(Y=1) = 1 - \operatorname{pr}(Y=0)$ is related to the unknown parameters through the relationship $\operatorname{pr}(Y=1) = \sum_{j,k,l} \operatorname{pr}(Y=1|G^o=j, G^m=k, X=l)\operatorname{pr}(X=l|G^m=k)\operatorname{pr}(G^o=j, G^m=k)$ by imposing conditional independence between G^o and X given G^m. In this model $T-1$ unknown parameters are involved for each possible G^m value, where T is the number of distinct values of $X_1, \ldots, X_n$ which can be large. In our method DEP, under the daLOG model, the total number of nuisance parameters related to the distribution $\operatorname{pr}(G^m, X)$ is reduced to $T-1$ plus the dimension of η. DEP-CLH maximized (3.1) using the same data as DEP. But, in the absence of an enumerated cohort, it specified a large number for N, say, 20k, and the total number of cases in the "cohort" as Nf.

In all simulation scenarios we also considered the standard prospective logistic regression analysis as implemented in the R function "glm" which is referred to as "LOGIT." LOGIT used the same design matrix as DEP. When the phenotype is on the health status of the offspring and no covariates are involved, the method implemented in the standalone software EMIM (version 3.22), which was designed for evaluating maternal, imprinting, and interaction genetics effects using multinomial modelling (Ainsworth et al. (2011a), Howey and Cordell (2012)), can also be used to analyze case-control mother–offspring pair data. Haplin, an R package, can also be used to detect the above effects based on a log-linear model (Gjessing and Lie (2006)). We note that EMIM and Haplin have comparable powers in detecting child and mother genetic effects through extensive simulations (Gjerdevik et al. (2019)). EMIM has been more often employed in practice, and it can be considerably faster than Haplin as the former was written in FORTRAN 77 while the latter was written in R. We therefore compared DEP with EMIM for testing main and interactive genetic effects, but EMIM cannot be applied to test gene-environment interactions due to lack of capacity to incorporate covariates.

In Section 3.1 results from a small-scale simulation study were presented for demonstrating the advantage of DEP. In Section 3.2 extensive simulation results were presented on the performance of DEP relative to IND and LOGIT. In Section 3.3 DEP was compared with EMIM in the absence of maternal covariates. In all simulation scenarios data were generated

under HWE (fixation index parameter $F = 0$) and each of the three modes of inheritance. HWE and correctly specified mode of inheritance and disease prevalence were implemented. Section 3.4 presented results on the robustness of DEP with respect to misspecification of the mode of inheritance, phenotype prevalence and deviation from HWE ($F > 0$).

3.1. *The performance of DEP relative to DEP-CLH.* We first compared DEP and DEP-CLH in terms of bias, statistical efficiency, robustness to mis-specification of model $\mathrm{pr}(X, G^m, G^c)$ and computation feasibility. Phenotype data Y was generated from model M5, where both genotypes were coded as the minor allele count divided by 2. We adjusted the intercept parameter in the penetrance model so that the population prevalence f was equal to 0.01 and fixed the MAF θ to be 0.2. A single environmental risk factor X was used. X was related to the maternal genotype G^m through either the daLOG model (2.3) with $\eta = \log(3.0)$ and $X \sim N(0, 1)$ or the linear model

$$X = \eta \times \{G^m - E(G^m)\} + e, \tag{3.2}$$

where the error term e is a standard normal random variable independent of G^m. We centered G^m at its mean and, consequently, X at zero in model (3.2) to reduce potential bias in the estimation of any considered method due to multicollinearity. Here, the linear model (3.2) guarantees that nonparametric specification of the conditional distribution of X, $\Pr(X|G^m)$, conforms with the model required by DEP-CLH but not with the daLOG model required by DEP. On the other hand, the daLOG model conforms with the conditional distribution of G^m, $\Pr(G^m|X)$, required by DEP but not with the model required by DEP-CLH. Our setup therefore permitted evaluation of robustness of the two methods with respect to misspecifying the relationship between X and G^m. The correlation between G^m and X characterized by η was set to be $\log(1.2)$ or $\log(3)$, mimicking moderate or strong association, respectively. The same penetrance model as that for generating case-control status Y was adopted. We generated a population of 10^7 mother–offspring pairs, from which 500 datasets with each consisting of $n_1 = 1000$ cases and $n_0 = 1000$ controls were selected. The true prevalence of Y, $f = 0.01$, was assumed known a priori and fixed in all subsequent analyses, except in the sensitivity study latter. To implement DEP-CLH, we set cohort size N artificially as 200,000 and specified the total numbers of ungenotyped case and control offspring as $M_1 = fN - n_1 = 1000$ and $M_0 = (1 - f)N - n_0 = 197{,}000$, respectively.

Among the 500 repetitions, DEP-CLH failed to converge for 13 datasets when data were generated under the linear model and three when data were generated under the daLOG model ($\eta = \log(3)$, nonzero genetic effects). DEP always converged and at a much faster rate. It took a laptop with a 2.29 GHz CPU no more than one second on average for DEP analyzing each simulated dataset, compared with 89 seconds for DEP-CLH. Table 3 and Table S1 (Zhang et al. (2020)) summarize the estimates. With moderate association between G^m and X (Table S1, Zhang et al. (2020)), the estimation biases of DEP were minor even if the G^m vs. X model was misspecified. DEP was comparable with DEP-CLH under model (2.3). With strong association (Table 3) the performance of DEP and DEP-CLH appeared to depend on the underlying model that generated data. Under the linear model DEP was slightly biased in regression parameter estimates, and the inflation of type-I error rates was minor. On the other hand, the empirical standard error of DEP-CLH ("SE") was visibly larger than the estimated one ("SEE") when the genetic effect was not zero, indicating that the algorithm for DEP-CLH might have difficulty in convergence in this situation. Under the daLOG model DEP had nearly unbiased estimates and well controlled type-I error rates, but DEP-CLH appeared to have large biases, inflated type-I error rates and sometimes lower power, especially for the interaction parameter. In all situations DEP had smaller standard errors. In summary, DEP was more efficient than DEP-CLH in terms of both increased computation speed and

reduced standard error. DEP was much more robust to misspecification of the distribution $\text{pr}(G^m, G^o, X)$. DEP-CLH may experience computation difficulty. We therefore did not consider DEP-CLH further in the simulation study.

3.2. *The performance of DEP relative to IND and LOGIT.* We then assessed the performance of DEP and IND with respect to bias, statistical efficiency, type-I error rate and power for testing genetic effects and assessing robustness with respect to misspecification of phenotype prevalence f. We also included results from LOGIT for comparison. Data were generated similarly as above, except that the daLOG model (2.3) was used to generate G^m and that all five penetrance models M1∼M5 were considered. The log-OR parameter for X, β_X, was fixed at $\log(1.5)$, and all the other parameters were set to be 0 under the null hypothesis or $\log(1.2)$ under the alternative hypothesis. We used the same coding for G^m in (2.3) as in the penetrance model, setting η equal to zero or $\log(3.0)$ corresponding to independence or strong correlation between G^m and X. We fixed phenotype prevalence f at 0.01 by adjusting the intercept parameter β_0 under each of the $2 \times 3 \times 2 \times 5 = 60$ parameter combinations. We analyzed each simulated dataset using all three methods.

Table 4 includes estimation bias and efficiency under the additive mode of inheritance, where G^m and G^o were independent of X. Results under the dominant and recessive modes of inheritance were similar (not shown). The estimation bias ("BIAS") for both DEP and IND was virtually zero, the mean standard error estimates ("SEE") was close to the empirical standard error ("SE") and the coverage probabilities of the 95% confidence intervals ("CP") were all close to the nominal level. DEP and IND had nearly the same efficiency for estimating genetic main effects, with asymptotic variance smaller than LOGIT by 20%∼ 36% across the five penetrance models. For estimating the interaction effect between G^m and X, $\beta_{G^m X}$, the asymptotic variance of IND, was smaller than that of DEP by 53% when G^o was not involved in the penetrance model (model M4). Under other models IND and DEP had nearly the same efficiency for estimating interaction parameters $\beta_{G^m X}$ and $\beta_{G^o X}$, although IND was slightly more efficient. Compared with LOGIT, IND and DEP had the largest efficiency gain for estimating $\beta_{G^o X}$, with variance reduced by 50%∼60%. Results were largely similar when log-ORs for all genetic effects were equal to zero (Table S2, Zhang et al. (2020)). When G^m and X were correlated, the averaged estimates by DEP remained close to the true values and had a much smaller variance than LOGIT, but IND appeared to be biased (Tables S3 and S4, Zhang et al. (2020)).

Table 5 presents type I error rate and power for testing the joint genetic effect ($H_0 : \boldsymbol{\beta}_G = 0$) under additive mode of inheritance at the 0.05 significance level. DEP and LOGIT maintained the nominal type-I error rate in all situations. But IND had dramatically inflated type-I error rates when G^m and X were correlated. DEP was more powerful than LOGIT in all situations, with the maximum power difference 26.7% when the combined effect of G^o and $G^o \times X$ was tested (model M3). When G^m and X were independent, IND was generally more powerful than DEP, with the maximum power difference being 23.8% under model M4 where the combined effects of G^m and $G^m \times X$ were tested. When the recessive or dominant mode of inheritance was applied (Tables S5 and S6, Zhang et al. (2020)), the power improvement of DEP over LOGIT became slightly larger, particularly for testing the main effect of G^o under model M1.

In the previous simulation studies the MAF θ was fixed at 0.2, and the sample size was 2000. We conducted additional simulations with a smaller MAF or a smaller sample size. Table S7 (Zhang et al. (2020)) displays type-I error rates and power for $\theta = 0.1$. The results are similar as before, except that DEP is slightly conservative (the minimum type-I error rate $= 0.029$). The conservative results might be due to slower weak convergence of the modified MLE when the MAF is small. Table S8 (Zhang et al. (2020)) displays some test results with a sample size of 400 (200 cases plus 200 controls), which exhibits a similar pattern as Table S7, except that the power became lower as expected.

TABLE 3

Estimation results for comparing DEP and DEP-CLH with strong association between maternal genotype and covariate (additive mode of inheritance, $f = 0.01$, $\theta = 0.2$, $\eta = \log(3)$)

			DEP[a]					DEP-CLH[b]				
$G^m \sim X^c$	log(OR)	True[d]	Bias[e]	SE[f]	SEE[g]	CP[h]	Power[i]	Bias[e]	SE[f]	SEE[g]	CP[h]	Power[i]
The linear model (3.2)	β_{G^o}	0.000	0.055	0.164	0.159	0.932	0.068	−0.021	0.166	0.157	0.939	0.061
	β_{G^m}	0.000	−0.069	0.167	0.161	0.924	0.076	0.027	0.179	0.172	0.933	0.067
	$\beta_{G^o X}$	0.000	−0.004	0.133	0.132	0.950	0.050	0.002	0.137	0.130	0.935	0.065
	$\beta_{G^m X}$	0.000	0.006	0.164	0.160	0.946	0.054	−0.003	0.177	0.167	0.949	0.051
	β_{G^o}	0.182	0.045	0.159	0.157	0.936	0.316	−0.030	0.161	0.154	0.909	0.183
	β_{G^m}	0.182	−0.087	0.158	0.157	0.910	0.108	0.018	0.175	0.171	0.945	0.237
	$\beta_{G^o X}$	0.000	−0.004	0.135	0.129	0.934	0.066	0.001	0.140	0.127	0.919	0.077
	$\beta_{G^m X}$	0.000	0.004	0.162	0.158	0.944	0.056	−0.007	0.187	0.165	0.939	0.057
	β_{G^o}	0.182	0.044	0.158	0.151	0.932	0.336	−0.035	0.202	0.151	0.932	0.203
	β_{G^m}	0.182	−0.074	0.146	0.152	0.924	0.098	0.036	0.169	0.166	0.940	0.242
	$\beta_{G^o X}$	−0.182	0.008	0.136	0.132	0.944	0.272	0.021	0.193	0.131	0.940	0.267
	$\beta_{G^m X}$	−0.182	0.004	0.158	0.161	0.956	0.198	−0.031	0.244	0.169	0.942	0.244

TABLE 3
(*Continued*)

			DEP[a]					DEP-CLH[b]				
$G^m \sim X^c$	log(OR)	True[d]	Bias[e]	SE[f]	SEE[g]	CP[h]	Power[i]	Bias[e]	SE[f]	SEE[g]	CP[h]	Power[i]
The daLOG model (2.3)	β_{G^o}	0.000	0.006	0.159	0.155	0.946	0.054	−0.064	0.166	0.153	0.923	0.077
	β_{G^m}	0.000	−0.003	0.165	0.159	0.942	0.058	0.057	0.187	0.166	0.919	0.081
	$\beta_{G^o X}$	0.000	−0.012	0.131	0.130	0.960	0.040	−0.009	0.135	0.129	0.960	0.040
	$\beta_{G^m X}$	0.000	0.012	0.161	0.162	0.946	0.054	0.107	0.169	0.160	0.909	0.091
	β_{G^o}	0.182	0.017	0.155	0.153	0.946	0.238	−0.066	0.174	0.151	0.909	0.147
	β_{G^m}	0.182	−0.001	0.163	0.156	0.936	0.216	0.074	0.193	0.164	0.911	0.382
	$\beta_{G^o X}$	0.000	−0.017	0.131	0.127	0.934	0.066	−0.011	0.132	0.126	0.932	0.068
	$\beta_{G^m X}$	0.000	0.000	0.160	0.159	0.950	0.050	0.103	0.159	0.159	0.895	0.105
	β_{G^o}	0.182	0.009	0.149	0.147	0.950	0.258	−0.064	0.145	0.145	0.930	0.124
	β_{G^m}	0.182	−0.017	0.150	0.151	0.952	0.208	0.057	0.159	0.158	0.926	0.321
	$\beta_{G^o X}$	−0.182	−0.022	0.130	0.129	0.946	0.358	−0.012	0.128	0.128	0.948	0.333
	$\beta_{G^m X}$	−0.182	0.025	0.162	0.162	0.938	0.164	0.116	0.159	0.160	0.902	0.078

[a] Our proposed method allowing for dependence between X and G^m; [b] the estimator proposed by Chen, Lin and Hochner (2012); [c] model relating G^m and X; [d] true value of the log-OR parameter; [e] difference between the mean estimate and true parameter value; [f] empirical standard error; [g] mean estimated standard error; [h] empirical coverage probability of 95% CI; [i] type-I error/power for testing genetic effect.

TABLE 4

Estimation results with nonzero genetic effects and unrelated G^m and X ($\beta_X = \log(1.5)$, $\beta_{G^o} = \beta_{G^m} = \beta_{G^m X} = \beta_{G^o X} = \log(1.2)$, $f = 0.01$, $\theta = 0.2$, $\eta = 0$)

		DEP[a]					IND[b]					LOGIT[c]			
Model	β	Bias[d]	SE[e]	SEE[f]	CP[g]	RE[h]	Bias[d]	SE[f]	SEE[h]	CP[g]	RE[h]	Bias[d]	SE[e]	SEE[f]	CP[g]
M1	β_{G^o}	0.000	0.137	0.138	0.951	1.35	0.000	0.137	0.138	0.950	1.35	0.003	0.159	0.160	0.948
	β_X	−0.001	0.046	0.047	0.950	1.00	−0.001	0.046	0.047	0.951	1.00	−0.001	0.046	0.047	0.950
M2	β_{G^m}	0.008	0.141	0.141	0.951	1.30	0.006	0.137	0.138	0.954	1.39	0.009	0.161	0.160	0.953
	β_X	0.001	0.046	0.047	0.948	1.00	0.001	0.046	0.047	0.948	1.00	0.001	0.046	0.047	0.948
M3	β_{G^o}	−0.002	0.145	0.146	0.951	1.30	−0.003	0.145	0.146	0.950	1.30	0.000	0.165	0.165	0.950
	β_X	0.005	0.053	0.053	0.954	1.17	0.003	0.052	0.053	0.952	1.21	0.003	0.057	0.058	0.954
	$\beta_{G^o X}$	−0.017	0.118	0.117	0.944	1.99	−0.014	0.111	0.109	0.945	2.25	−0.004	0.166	0.166	0.949
M4	β_{G^m}	−0.007	0.147	0.147	0.949	1.25	−0.008	0.146	0.147	0.949	1.28	−0.005	0.165	0.165	0.947
	β_X	0.004	0.058	0.057	0.947	1.04	0.003	0.053	0.053	0.948	1.23	0.003	0.059	0.058	0.948
	$\beta_{G^m X}$	0.008	0.158	0.157	0.946	1.14	0.011	0.109	0.110	0.956	2.40	0.012	0.169	0.166	0.948
M5	β_{G^o}	−0.014	0.154	0.154	0.955	1.54	−0.015	0.154	0.154	0.955	1.54	−0.016	0.191	0.193	0.955
	β_{G^m}	0.018	0.153	0.153	0.947	1.56	0.017	0.153	0.153	0.947	1.56	0.019	0.191	0.193	0.949
	β_X	−0.002	0.059	0.059	0.954	1.06	−0.002	0.055	0.055	0.950	1.23	−0.003	0.061	0.062	0.955
	$\beta_{G^o X}$	0.002	0.125	0.126	0.953	2.50	0.002	0.125	0.126	0.953	2.50	0.012	0.197	0.194	0.947
	$\beta_{G^m X}$	−0.002	0.168	0.169	0.953	1.39	−0.001	0.126	0.126	0.951	1.46	0.000	0.198	0.194	0.946

[a] Our proposed method allowing for dependence between X and G^m; [b] our proposed method with independence assumption between X and G^m; [c] the conventional logistic regression method; [d] difference between the mean estimate and true parameter value; [e] empirical standard error; [f] mean estimated standard error; [g] empirical coverage probability of the 95% confidence interval; [h] relative efficiency defined as the asymptotic variance of LOGIT divided by that of the target method.

TABLE 5

Type-I error rates and powers for testing the joint genetic effect under the additive mode of inheritance ($f = 0.01$, $\theta = 0.2$). Under the null hypothesis, $\beta_{G^o} = \beta_{G^m} = \beta_{G^m X} = \beta_{G^o X} = 0$; under the alternative hypothesis, $\beta_{G^o} = \beta_{G^m} = \beta_{G^m X} = \beta_{G^o X} = \log(1.2)$. The log-OR for X was $\beta_X = \log(1.5)$

		Null hypothesis			Alternative hypothesis		
η^d	Model	DEP[a]	IND[b]	LOGIT[c]	DEP[a]	IND[b]	LOGIT[c]
0	M1	0.054	0.054	0.056	0.258	0.260	0.208
	M2	0.051	0.049	0.048	0.276	0.282	0.232
	M3	0.048	0.050	0.045	0.541	0.582	0.319
	M4	0.049	0.048	0.048	0.403	0.641	0.341
	M5	0.056	0.060	0.050	0.831	0.954	0.713
log(3.0)	M1	0.049	0.059	0.047	0.286	–[e]	0.210
	M2	0.048	0.948	0.046	0.284	–	0.235
	M3	0.053	0.997	0.049	0.648	–	0.381
	M4	0.050	1.000	0.050	0.525	–	0.465
	M5	0.050	1.000	0.048	0.900	–	0.818

[a]Our proposed method allowing for dependence between X and G^m; [b]our proposed method with independence assumption between X and G^m; [c]the conventional logistic regression method. [d]The maternal genotype was either independent of ($\eta = 0$) or strongly correlated with the environmental risk factor ($\eta = \log(3.0)$), as specified in the daLOG model. [e]Power not displayed because of inflated type-I error rates.

3.3. *The performance of DEP relative to EMIM.* EMIM can be used to infer genetic effects on offspring phenotypes using case-control mother–offspring data, but it cannot handle maternal covariates. In both EMIM and DEP we considered testing single main genetic effects and joint genetic effect. The two methods differed in the joint genetic effect: EMIM tested two main genetic effects jointly, while DEP involved two additional gene-envirnoment interaction effects. In both methods we incorporated the HWE and additive mode of inheritance (the same as the multiplicative model for EMIM) for both maternal and offspring genetic effects. Data were generated according to the penetrance model (2.1) and the daLOG model (2.3). A single covariate X following the standard normal distribution was considered, and η was set at zero in the daLOG model because, otherwise, EMIM would have inflated type-I error rates due to the confounding effect of X. Phenotype prevalence f was set at 0.01, the MAF at 0.2 and the covariate effect at $\beta_X = \log(1.5)$. Three parameter combinations were considered:

M00: $\beta_{G^o} = \beta_{G^m} = \beta_{G^o X} = \beta_{G^m X} = 0$;
M10: $\beta_{G^o} = \beta_{G^m} = \log(1.2)$, $\beta_{G^o X} = \beta_{G^m X} = 0$;
M11: $\beta_{G^o} = \beta_{G^m} = \log(1.2)$, $\beta_{G^o X} = \beta_{G^m X} = -\log(1.2)$.

For each parameter combination, 1000 case pairs and 1000 control pairs were generated, and the hypothesis tests were carried out for the two main genetic effects and joint genetic effect. Model M00 allowed assessment of type-I error rates, and models M10 and M11 allows power comparison in the presence and absence of interaction effects. All tests were performed at 0.05 significance level, and the simulation was repeated 5000 times (Table S9, Zhang et al. (2020)).

The type-I error rates appeared to be well controlled at the nominal level for both methods. In the absence of gene-enviroment interaction effects (M10), EMIM could be much more powerful than DEP when testing joint genetic effect. This was expected since the former did not incorporate the covariate X, thus no gene-environment interation was considered, but DEP involved two more interaction terms with zero effect. The power of the two methods was

similar when single main genetic effects were tested. On the other hand, DEP could be much more powerful in testing joint genetic effect in the presence of interaction effects (models M11).

3.4. *Robustness of DEP with respect to misspecification of the fixation index parameter, diesease prevalence, or mode of inheritance.* First, we assessed the robustness of DEP and IND by misspecifying the disease prevalence f. The true f value was low (0.01), but was misspecified to be much larger (0.2), or was moderately high (0.2), but was misspecified to be much smaller (0.01). We considered various association strength between G^m and X ($\eta = \log(1.2)$ or $\log(3.0)$). The underlying parameters and the corresponding estimation and test results are presented in Table 6 and Table S10 (Zhang et al. (2020)). Interestingly, DEP appeared to be unbiased except for the intercept parameter. The type-I error rates and power were minimally affected (Table 6) unless the misspecification of f was severe (Table S10, Zhang et al. (2020)). As shown in Table S10, the power of DEP with the f value biased downward appeared to be more powerful than that with the f value biased upward, regardless of the magnitude of the true f value. It was intriguing that IND was much less powerful than DEP in spite of inflated type-I error rates (data not shown). For example, under model M5 the bias was 0.197 by IND, compared with 0.014 by DEP and the corresponding power was 0.454 and 0.581, respectively. We consistently observed the inferior power of IND under different parameter values.

Second, we assess the robustness of DEP with respect to HWE for maternal genotypes. We generated data with nonzero fixation index parameter F (0.05, 0.1 or 0.2) so that HWE did not hold for maternal genotypes but implemented DEP under HWE. The remaining parameters were identical to those for Table 6. The misspecification of F appeared to have negligible impact on the estimation and testing results for all regression parameters (Table S11, Zhang et al. (2020)).

Lastly, we studied the impact of misspecifying the mode of inheritance on DEP. We generated data under one mode of inheritance but implemented DEP under a different mode of inheritance in each simulation scenario. The remaining parameters were again identical to those for Table 6. Using misspecified mode of inheritance did not lead to inflation in the type-I error rates but could result in considerable loss in power (Table S12, Zhang et al. (2020)). The additive mode of inheritance appeared to be the most robust in terms of the extent of power loss.

4. Analysis of genetic association studies of GDM and birth weight. We analyzed data from two genetic association studies, one on gestational diabetes mellitus which has a maternal phenotype, and the other on infant birth weight which has an offspring phenotype. Both studies have paired mother–offspring genotype data and maternal covariate oberservations. We applied all four methods, DEP, DEP-CLH, IND and LOGIT to fit penetrance model M5 to analyze both studies under the additive mode of inheritance. Because it was designed for offspring phenotypes, we also used EMIM to analyze the second study. High prepregnancy BMI (pp-BMI) has been established to be associated with both maternal pregnancy health and low birth weight (Abubakari, Kynast-Wolf and Jahn (2015), Frederick et al. (2008), Mallia et al. (2017)). For both phenotypes we assessed selected candidate genes through joint testing of their main effects and interactions with pp-BMI, as such joint tests may be more powerful than testing the marginal genetic effect (Kraft et al. (2007)). For significant SNPs we were also interested in testing whether they modified the effect of pp-BMI. In both studies it was reasonable to consider pp-BMI as independent of the offspring genotype conditional on the maternal genotype. All SNPs that we assessed were common with MAF above 0.05.

TABLE 6

Simulation results with moderatly misspecified prevalence ($\theta = 0.2$, $\beta_{G^o} = \beta_{G^m} = \log(1.2)$ and $\beta_{G^oX} = \beta_{G^mX} = -\log(1.2)$). The maternal genotype was moderately correlated with the environmental risk factor ($\eta = \log(1.2)$ in the daLOG model). The phenotype prevalence f was specified either at the true value 0.01 or misspecified as 0.02. The other parameters were the same as those in Table 5

			DEP[a]		DEP[a,d]		IND[b]		IND[a,d]		LOGIT[c]	
Model	Log-OR	True[e]	Bias[f]	Power[g]	Bias[f]	Power[g]	Bias[f]	Power[g]	Bias[f]	Power[g]	Bias[f]	Power[g]
M3	β_0	−4.699	0.001	0.315	0.704	0.316	0.005	0.180	0.708	0.181	4.596	0.236
	β_{G^o}	0.182	−0.012		−0.011		−0.033		−0.032		−0.007	
	β_X	0.405	0.002		0.003		−0.016		−0.016		0.003	
	β_{G^oX}	0.182	0.004		0.002		0.093		0.092		0.001	
M4	β_0	−4.698	0.000	0.251	0.703	0.250	−0.001	0.196	0.702	0.193	4.595	0.217
	β_{G^m}	0.182	−0.007		−0.007		−0.011		−0.011		−0.008	
	β_X	0.405	0.005		0.005		−0.033		−0.033		0.004	
	β_{G^mX}	−0.182	0.007		0.006		0.191		0.190		0.010	
M5	β_0	−4.725	0.001	0.581	0.705	0.581	0.001	0.454	0.704	0.454	4.596	0.495
	β_{G^o}	0.182	−0.018		−0.018		−0.021		−0.022		−0.013	
	β_{G^m}	0.182	0.007		0.007		0.004		0.005		0.009	
	β_X	0.405	0.000		0.000		−0.038		−0.037		0.000	
	β_{G^oX}	−0.182	0.000		−0.001		0.001		−0.001		0.002	
	β_{G^mX}	−0.182	0.014		0.015		0.197		0.196		0.012	

[a] Our proposed method allowing for dependence between X and G^m; [b] our proposed method with independence assumption between X and G^m; [c] the conventional logistic regression method; [d] the true prevalence was 0.01 but mis-specified as 0.02; [e] true parameter value; [f] difference between the mean estimate and true parameter value; [g] power for testing the joint genetic effect.

4.1. *The gestational diabetes mellitus study.* We analyzed 1156 mother–offspring pairs from Tianjin Gestational Diabetes Mellitus Prevention Program (Hu et al. (2012)), where 578 mothers had GDM and 578 were healthy during pregnancy. Maternal information, including sociodemographic characteristics, pregnancy outcomes, and lifestyle in the past year was collected by a selfadministered questionnaire. The main goal of the current analysis was to assess whether the 15 SNPs that have been reported to predispose GDM (Radha et al. (2016)) were also associated with the risk of GDM in the current study and whether the association was driven by maternal or offspring genes. We analyed 14 SNPs because one had an MAF smaller than 0.05, and used GDM prevalence 14% as observed in Tianjin women. Results were nearly identical when lower (8%) or higher (20%) values of prevalence were used (unreported). A p-value less than 0.05/14 was considered statistically significant which accounted for multiplicity in testing 14 SNPs.

Table S13 (Zhang et al. (2020)) displays the estimated MAFs, p-values for testing HWE and p-values for testing independence between pp-BMI and maternal genotypes. All MAFs were greater than 0.1. None of the SNPs showed evidence of deviation from HWE, therefore, we incorporated HWE in all the analysis where applicable. The p-values for testing the joint genetic effects are displayed in Figure 1. SNP rs2237895, which is in the voltage-gated potassium channel gene, was found to be statistically significant, and the effect appeared to be mainly driven by the maternal SNP, as confirmed by the marginal analysis of maternal and offspring SNPs (OR 1.89; 95% CI [1.38, 2.62]). Two other SNPs, rs1387153 in gene Melatonin receptor 1B and rs4406920 in gene Insulin-like Growth Factor 2 had p-values less than 0.05 but did not reach significance after Bonferroni correction. The results suggested that both maternal SNPs may be associated with the risk of GDM, and offspring SNPs may also be associated but through modifying the effect of pp-BMI. These results may be worth exploring in studies of larger sizes. The estimated OR parameters for these three SNPs are included in Table 7. IND yielded the three smallest p-values which was reasonable since pp-BMI appeared to be uncorrelated with the three corresponding maternal genotypes. Tables S14∼S17 (Zhang et al. (2020)) include the estimation and test results of the main effects and interactions effects for all SNPs.

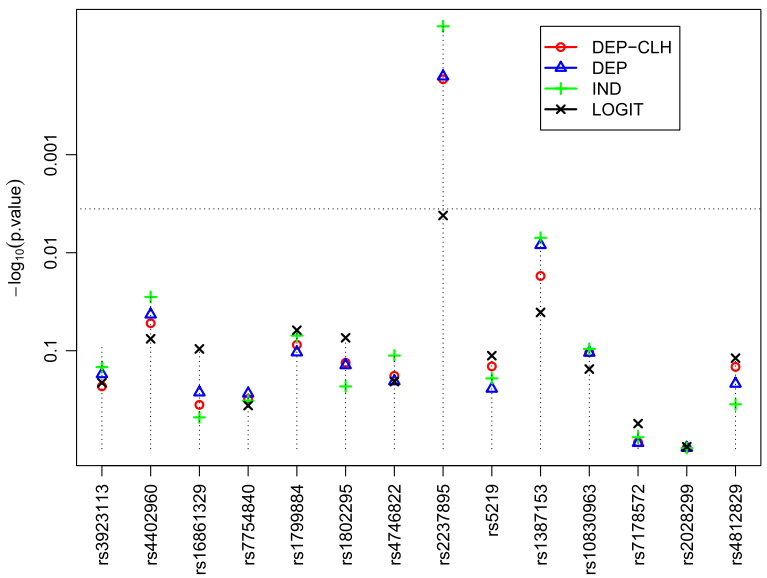

FIG. 1. *P-values for testing the joint effect of maternal and offspring genotypes on the risk of GDM. DEP-CLH, the method proposed by Chen, Lin and Hochner (2012); DEP, our proposed method allowing for dependence between X and G^m; IND, our proposed method with independence assumption between X and G^m; LOGIT, the standard logistic regression method.*

TABLE 7

Genetic effect estimates for the three SNPs that were significantly (nominal level = 0.05 without Bonferroni correction) associated with the risk of GDM in the Tianjin Gestational Diabetes Mellitus Prevention Program

SNP		DEP[a]			IND[b]			LOGIT[c]		
rs2237895	Log-OR	$\widehat{\text{Log-OR}}$[d]	SE[e]	P-value[f]	$\widehat{\text{Log-OR}}$[d]	SE[e]	P-value[f]	$\widehat{\text{Log-OR}}$[d]	SE[e]	P-value[f]
	β_{G^o}	−0.04	0.184	0.835	−0.04	0.183	0.851	0.06	0.236	0.809
	β_{G^m}	0.45	0.181	*0.013*	0.45	0.180	*0.013*	0.45	0.238	0.059
	β_{G^oX}	0.30	0.149	*0.042*	0.30	0.150	*0.043*	0.43	0.249	0.084
	β_{G^mX}	−0.12	0.226	0.603	−0.06	0.150	0.695	−0.16	0.240	0.493
	P-joint[g]		*0.043*			*0.028*			0.052	
rs2237895	Log-OR	$\widehat{\text{Log-OR}}$[a]	SE[b]	P-value[c]	$\widehat{\text{Log-OR}}$[a]	SE	P-value	$\widehat{\text{Log-OR}}$[a]	SE	P-value
	β_{G^o}	−0.41	0.185	*0.028*	−0.41	0.184	*0.025*	−0.42	0.228	0.063
	β_{G^m}	0.77	0.170	**6.4e−6**	0.76	0.170	**7.0e−6**	0.83	0.228	**2.8e−4**
	β_{G^oX}	0.14	0.136	0.291	0.14	0.136	0.292	0.28	0.249	0.269
	β_{G^mX}	−0.20	0.211	0.340	−0.29	0.136	0.034	−0.29	0.234	0.222
	P-joint		**1.6e−4**			**4.9e−5**			**4.2e−3**	
rs1387153	Log-OR	$\widehat{\text{Log-OR}}$[a]	SE[b]	P-value[c]	$\widehat{\text{Log-OR}}$[a]	SE	P-value	$\widehat{\text{Log-OR}}$[a]	SE	P-value
	β_{G^o}	−0.02	0.165	0.882	−0.03	0.164	0.859	0.15	0.222	0.489
	β_{G^m}	0.43	0.165	*0.009*	0.43	0.164	*0.010*	0.25	0.217	0.252
	β_{G^oX}	0.31	0.129	*0.016*	0.31	0.129	*0.016*	0.58	0.233	*0.013*
	β_{G^mX}	0.05	0.197	0.807	−0.031	0.14	0.819	−0.07	0.222	0.743
	P-joint		*0.008*			*0.007*			*0.017*	

[a] Our proposed method allowing for dependence between X and G^m; [b] our proposed method with independence assumption between X and G^m; [c] the conventional logistic regression method; [d] estimated log-OR parameters; [e] asymptotic standard error; [f] p-value of the Wald test for each single effect; [g] p-value of the Wald test for the joint genetic effect.

4.2. *The Jerusalem perinatal study.* We analyzed a case-control mother–offspring pair study nested within the JPS (Harlap et al. (2007)) to assess the association between gene PPARGC1A and low birth weight. The JPS was a prospective cohort study that included 17,003 births to residents of Jerusalem between years 1974 to 1976. Between years 2007 and 2009, a subsample of 1250 offspring in the JPS and their mothers were interviewed and genotyped in a follow-up study for assessing cardio-metabolic traits (Hochner et al. (2012)). Offspring whose birth weight was below 2500 grams or above 4000 grams or whose mothers' pp-BMI was above 27 were over-represented in this subsample. We analyzed a subset of 691 pairs with pp-BMI less than 25, consisting of 125 pairs with low birth weight ($<$2500 grams) and 566 control pairs with normal or high birth weight ($>$2500 grams). The gene PPARGC1A encodes a protein that regulates genes involved in energy metabolism and, therefore, may be associated with birth weight and could modify the effect of pp-BMI. We considered both offspring and maternal genes, as both may be implicated in the risk of low birth weight (Lunde et al. (2007), Infanterivard (2007)). The additive mode of inheritance was assumed for all the SNPs. Among the 8238 eligible offspring in the JPS cohort (mothers' pp-BMI $<$ 25), 297 had low birth weight. Therefore, we estimated the prevalence of low birth weight as $297/8238 \approx 0.036$. A significance level of $0.05/24 \approx 0.002$ instead of 0.05 was considered statistically significant in order to account for multiplicility in testing 24 SNPs. Table S18 (Zhang et al. (2020)) displays the estimated MAFs, p-values for testing HWE, p-values for testing joint genetic association and p-values for testing independence between pp-BMI and maternal genotypes. All MAFs were greater than 0.1. None of the SNPs had a p-value less than 0.05 for testing HWE; therefore, we incorporated HWE in all methods where applicable.

The p-values for testing the joint genetic effect are displayed in Figure 2, where EMIM had two degrees of freedom because the interaction term was not able to be incorporated and the other tests all had four degrees of freedom. Except for one SNP, the p-value of DEP was smaller than or similar to that of DEP-CLH, suggesting higher efficiency of DEP. Two SNPs were found significant by at least one method. SNP rs3774921 was found significant

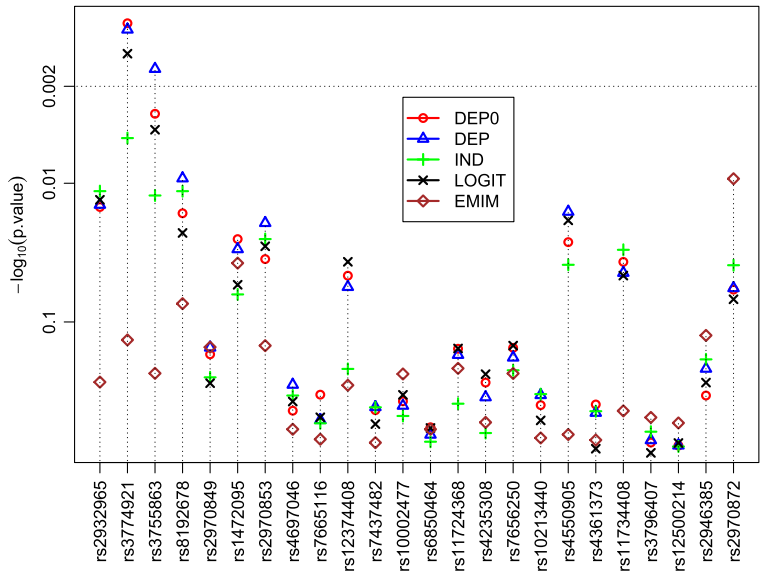

FIG. 2. *P-values for testing the joint effect of maternal and offspring genes PPARGC*1*A* (24 *SNPs*) *on the risk of low birth weight in the Jerusalem Perinatal Study. DEP-CLH, the method proposed by Chen, Lin and Hochner* (2012); *DEP, our proposed method allowing for dependence between X and* G^m; *IND, our proposed method with independence assumption between X and* G^m; *LOGIT, the standard logistic regression method*; *EMIM, the method developed by Ainsworth et al.* (2011*a*) *and Howey and Cordell* (2012).

by DEP, DEP-CLH and LOGIT but not by IND and EMIM, with corresponding p-values 0.0008, 0.002, 0.001 and 0.005, respectively. SNP rs3755863 was found significant by DEP (p-value 0.001) but not by the other methods. The detailed results of DEP, IND, and LOGIT for the two SNPs are presented in Table 8. The significance of the joint genetic effect for both SNPs appeared to be driven by their interactions with pp-BMI. In fact, the p-value by DEP for testing the null hypothesis of no interaction with maternal SNPs, $\beta_{G^m X} = 0$, was 0.002 for SNP rs3774921 and <0.001 for SNP rs3755863. Interestingly, the direction of this interaction effect was opposite for the two SNPs, with corresponding log-ORs equal to 1.04 and -1.19, respectively. DEP and LOGIT yielded similar interaction test p-values, and IND yielded much larger p-values for both SNPs. The estimated interaction effect by IND was weaker for both SNPs. The interaction log-OR, $\hat{\beta}_{G^m X}$, was estimated to be 1.04 by DEP and 0.79 by IND for SNP rs3774921 and -1.19 by DEP and -0.91 by IND for SNP rs3755863. The apparent difference can be explained by the significant correlation between G^m and pp-BMI, as the parameter η in the daLOG model for the two SNPs was estimated to be -0.149 (p-value $= 0.023$) and 0.166 (p-value $= 0.004$), respectively (Table S18, Zhang et al. (2020)). The estimate of the maternal genotype main effect β_{G^m} by LOGIT differed from those by DEP and IND, but the three corresponding confidence intervals were wide and largely overlapped. Results for the remaining SNPs are included in Tables S19∼S22 (Zhang et al. (2020)). The p-value of DEP tended to be smaller than that of IND as the association between pp-BMI and maternal genotype became stronger (Figure S1, Zhang et al. (2020)). There was no evidence that any of the offspring SNPs modified the effect of pp-BMI (Table S21, Zhang et al. (2020)).

5. Discussion. The case-control mother–offspring pair design has been popularly used for identifying genes that are associated with obstetrical and early-life phenotypes. It allows investigation on joint effects of both maternal and offspring genes and environmental risk factors. Our new method overcomes the computation difficulty of the Chen, Lin and Hochner (2012) method and is fast enough to allow genomewide investigations. It is, generally, statistically more efficient than the other considered methods as shown in simulation studies. A key to success of our method is a novel regression model relating the maternal genotype and environmental factors which makes it plausible to obtain closed forms for the limiting Lagrange multipliers involved in derivation of the profile likelihood function. The resultant modified profile likelihood function is generally concave with unknown parameters so that its maximizer can be numerically obtained using any optimization algorithm designed for finding global or local maximizers. Our modified profile likelihood method is more efficient than the standard prospective logistic regression analysis and was able to identify the genetic effects that other methods would have missed in the analysis of the real data examples. We also considered a variant of our method that additionally requires independence between environmental risk factors and maternal genotypes, in the same spirit as the literature on exploiting gene-environment independence to increase efficiency for making inference on gene-environment interaction effects (Chatterjee and Carroll (2005), Piegorsch, Weinberg and Taylor (1994), Umbach and Weinberg (1997)). The estimation bias and inflation in the type-I error rate could be severe should the independence constraint be violated. Enforcing the additional constraint of independence led to meaningful efficiency gain for assessing the interaction effect between a maternal SNP and an environmental factor. But the gain was largely marginal except for under the simplest model involving only the main and interaction effects of the two variables. The conditional independence requirement most likely holds true in the setting of obstetrical and early-life outcomes. Therefore, our method is attractive because of its robustness and high efficiency. For analyzing offspring phenotypes, in the absence of confounding covariates and gene-environment interactions, existing methods as

TABLE 8

Genetic effect estimates for the two SNPs that were significantly associated with the risk of low birth weight in the Jerusalem perinatal study

SNP		DEP[a]			IND[b]			LOGIT[c]		
rs3774921	Log-OR	$\widehat{\text{Log-OR}}$[d]	SE[e]	P-value[f]	$\widehat{\text{Log-OR}}$[d]	SE[e]	P-value[f]	$\widehat{\text{Log-OR}}$[d]	SE[e]	P-value[f]
	β_{G^o}	0.68	0.323	0.036	0.68	0.323	0.035	0.65	0.343	0.058
	β_{G^m}	−0.23	0.347	0.510	−0.23	0.345	0.507	−0.17	0.349	0.635
	β_{G^oX}	0.25	0.336	0.463	0.25	0.334	0.461	0.17	0.340	0.615
	β_{G^mX}	1.04	0.338	**1.98E−3**	0.79	0.315	0.012	1.06	0.355	2.84E−3
	P-joint[d]			**7.74E−4**			4.67E−3			**1.16E−3**
rs3755863	Log-OR	$\widehat{\text{Log-OR}}$[a]	SE[b]	P-value[c]	$\widehat{\text{Log-OR}}$[a]	SE	P-value	$\widehat{\text{Log-OR}}$[a]	SE	P-value
	β_{G^o}	−0.46	0.304	0.127	−0.47	0.304	0.124	−0.41	0.337	0.223
	β_{G^m}	−0.05	0.320	0.884	−0.04	0.320	0.889	−0.21	0.345	0.539
	β_{G^oX}	0.34	0.323	0.295	0.33	0.320	0.297	0.27	0.339	0.426
	β_{G^mX}	−1.19	0.345	**5.44E−4**	−0.91	0.333	6.31E−3	−1.19	0.339	**4.25E−4**
	P-joint			**1.45E−3**			0.012			4.11E−3

[a] Our proposed method allowing for dependence between X and G^m; [b] our proposed method with independence assumption between X and G^m; [c] the conventional logistic regression method; [d] estimated log-OR parameters; [e] asymptotic standard error; [f] p-value of the Wald test for each single effect; [g] p-value of the Wald test for the joint genetic effect.

implemented in EMIM (Ainsworth et al. (2011a), Howey and Cordell (2012)) can be more powerful for testing genetic effects.

We note that, under model (2.3), maternal and offspring genotypes may have different distributions. When G^m is independent of X as indicated by $\eta = 0$, the marginal distribution of G^m is the same as that of the offspring genotype G^o. With nonzero η, the two marginal distributions are not the same any more (see Table S23 in Appendix (Zhang et al. (2020)) for some examples). However, in our simulation studies presented in Section 3, the estimation bias based on the proposed modified profile likelihood method appeared to be minor even when the correlation between G^m and X was strong. Under the daLOG model the distributions of maternal and offspring genotypes were somewhat different, as mentioned previously. As a result, DEP-CLH was biased, as it requires a linear model between X and G^m, which is not satisfied under the daLOG model.

Our method requires known phenotype prevalence. Fortunately, its misspecification, while leading to biased estimates of the intercept parameter β_0 in models M1$\sim$M5, appeared to have negligible impact on estimation and testing of genetic effects unless the misspecification is severe. This aligns with the well-known result that incorporating known disease prevalence only has impact on estimating the intercept parameter in standard logistic regression analysis. We leave theoretical exploration of this observation in broader scenarios to future work. As shown in Table S10, the power of testing genetic association is higher when the bias in the specified prevalence is downward than that when the bias is upward. Therefore, we suggest that a conservative value of prevalence be used when applying our method for analysis.

Under the case-control mother–offspring pair design, it is feasible to assess parent-of-origin effects owing to the pairing of mother–offspring genotype data. Parent-of-origin effects of the APOB gene were recently reported for adiposity in early adulthood (Hochner et al. (2012)). With mother–offspring paired genotype data, the parental source of the two offspring alleles is unknown when both genotypes are heterozygous. Statistical methods are available for accommodating ambiguity in parental sources (Ainsworth et al. (2011b), Shi et al. (2008), Weinberg and Umbach (2005), Weinberg, Wilcox and Lie (1998), Yang and Lin (2013)). But none of them allows incorporation of covariates, which is nontrivial due to the retrospective case-control sampling and incomplete information on parental origins of offspring alleles. For offspring phenotypes it would be very interesting to extend our method to accommodate X-chromosome data. In addition, it would be interesting to extend the method for studying haplotype effects, where haplotype phases can be inferred from multimarker genotype data for both the mother and offspring. We will pursue this topic in future work.

Our dyad design has been specifically adopted for studying obstetrical complications which are maternal phenotypes. In this sense the dyad design is unique in its own way. For studying offspring phenotypes, failure to consider paternal information in the mother–offspring paired design may lead to paternal confounding bias. Readers are referred to Umbach and Weinberg (2000) and Kistner, Shi and Weinberg (2009) for more discussions. In practice, it is difficult to imagine a situation that one would intentionally design such a dyad study for studying offspring phenotypes, and at least attempts would be made to obtain triads. Just as in the example JPS study, one scenario that a study of offspring phenotypes would include exclusively dyads would be a secondary analysis of a study on maternal phenotypes, where paternal information may not have been collected. When both case-parent and control-parent triad data are available, our current method can be applied for analysis with paternal genotype data ignored. But paternal genotype data can be incorporated by straightforward extension. It can contribute to more precise estimation of the minor allele or haplotype frequencies (Wilcox, Weinberg and Lie (1998)), thereby leading to increased power for testing genetic associations and alternative hypotheses such as maternal-fetal genotype interactions. We will explore the merit of such extension relative to existing case- and control-triad methods in future work (Dudbridge (2008), Gjessing and Lie (2006), Skare et al. (2012)).

When genetic data are available only for case-parent trios, methods are available for testing gene-environent interaction effects (e.g., Shi, Umbach and Weinberg (2011)). However, methods for testing gene-environment interaction effects are largely unexplored when paternal genetic data is not available. We will pursue this issue in the future.

Acknowledgments. We thank the editor, associate editor and three reviewers for their insightful comments.

The first author was supported in part by NIH grants R21-ES020811 and R01-ES016626 and NSFC grant 11771096.

The second author was supported in part by NIH grants R21-ES020811.

The sixth author was supported in part by NIH grants R21-ES020811 and R01-ES016626.

SUPPLEMENTARY MATERIAL

Supplement to "An efficient and computationally robust statistical method for analyzing case-control mother–offspring pair genetic association studies": (Figure S1, Tables S1–S23, derivation of the log profile likelihood function (2.5) and equations (2.8), and proof of the asymptotic equivalence ofprofile MLE and modified profile MLE mentioned in Section 2.3.

REFERENCES

ABUBAKARI, A., KYNAST-WOLF, G. and JAHN, A. (2015). Maternal determinants of birth weight in northern Ghana. *PLoS ONE* **10** e0135641. https://doi.org/10.1371/journal.pone.0135641

AGRESTI, A. (2013). *Categorical Data Analysis*, 3rd ed. *Wiley Series in Probability and Statistics*. Wiley- Interscience, Hoboken, NJ. MR3087436

AINSWORTH, H. F., UNWIN, J., JAMISON, D. L. and CORDELL, H. J. (2011a). Investigation of maternal effects,

maternal-foetal interactions and parent-of-origin effects (imprinting), using mothers and their offspring. *Genet. Epidemiol.* **35** 19–45.

AINSWORTH, H. F., UNWIN, J., JAMISON, D. L. and CORDELL, H. J. (2011b). Investigation of maternal effects, maternal-fetal interactions and parent-of-origin effects (imprinting), using mothers and their offspring. *Genet. Epidemiol.* **35** 19–45. https://doi.org/10.1002/gepi.20547

CHATTERJEE, N. and CARROLL, R. J. (2005). Semiparametric maximum likelihood estimation exploiting gene-environment independence in case-control studies. *Biometrika* **92** 399–418. MR2201367 https://doi.org/10.1093/biomet/92.2.399

CHEN, J., LIN, D. and HOCHNER, H. (2012). Semiparametric maximum likelihood methods for analyzing genetic and environment effects with case-control mother-child pair data. *Biometrics* **68** 869–877. MR3055191 https://doi.org/10.1111/j.1541-0420.2011.01728.x

CHEN, J., ZHENG, H. and WILSON, M. L. (2009). Likelihood ratio tests for maternal and fetal genetic effects on obstetric complications. *Genet. Epidemiol.* **33** 526–538.

DUDBRIDGE, F. (2008). Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. *Hum. Hered.* **66** 87–98. https://doi.org/10.1159/000119108

FREDERICK, I. O., WILLIAMS, M. A., SALES, A. E., MARTIN, D. P. and KILLIEN, M. (2008). Pre-pregnancy body mass index, gestational weight gain, and other maternal characteristics in relation to infant birth weight. *Matern. Child Health J.* **12** 557–567.

GJERDEVIK, M., JUGESSUR, A., HAALAND, Ø. A., ROMANOWSKA, J., LIE, R. T., CORDELL, H. J. and GJESSING, H. K. (2019). Haplin power analysis: A software module for power and sample size calculations in genetic association analyses of family triads and unrelated controls. *BMC Bioinform.* **20** 165. https://doi.org/10.1186/s12859-019-2727-3

GJESSING, H. K. and LIE, R. T. (2006). Case-parent triads: Estimating single-and double-dose effects of fetal and maternal disease gene haplotypes. *Ann. Hum. Genet.* **70** 382–396.

GODDARD, K. A., TROMP, G., ROMERO, R., OLSON, J. M., LU, Q., XU, Z., PARIMI, N., NIEN, J. K., GOMEZ, R. et al. (2007). Candidate-gene association study of mothers with pre-eclampsia, and their infants, analyzing 775 SNPs in 190 genes. *Hum. Hered.* **63** 1–16.

HARLAP, S., DAVIES, A. M., DEUTSCH, L., CALDERON-MARGALIT, R., MANOR, O., PALTIEL, O., TIRAM, E., YANETZ, R., PERRIN, M. C. et al. (2007). The Jerusalem perinatal study cohort, 1964–2005: Methods and a review of the main results. *Paediatr. Perinat. Epidemiol.* **21** 256–273.

HOCHNER, H., FRIEDLANDER, Y., CALDERON-MARGALIT, R., MEINER, V., SAGY, Y., AVGIL-TSADOK, M., BURGER, A., SAVITSKY, B., SISCOVICK, D. S. et al. (2012). Associations of maternal prepregnancy body mass index and gestational weight gain with adult offspring cardiometabolic risk factors: The Jerusalem perinatal family follow-up study. *Circulation* **125** 1381–1389.

HOWEY, R. and CORDELL, H. J. (2012). PREMIM and EMIM: Tools for estimation of maternal, imprinting and interaction effects using multinomial modelling. *BMC Bioinform.* **13** 149. https://doi.org/10.1186/1471-2105-13-149

HU, G., TIAN, H., ZHANG, F., LIU, H., ZHANG, C., ZHANG, S., WANG, L., LIU, G., YU, Z. et al. (2012). Tianjin gestational diabetes mellitus prevention program: Study design, methods, and 1-year interim report on the feasibility of lifestyle intervention program. *Diabetes Res. Clin. Pract.* **98** 508–517.

INFANTERIVARD, C. (2007). Studying genetic predisposition among small-for-gestational-age newborns. *Semin. Perinatol.* **31** 213–218.

KANAYAMA, N., TAKAHASHI, K., MATSUURA, T., SUGIMURA, M., KOBAYASHI, T., MONIWA, N., TOMITA, M. and NAKAYAMA, K. (2002). Deficiency in p57Kip2 expression induces preeclampsia-like symptoms in mice. *Mol. Hum. Reprod.* **8** 1129–1135.

KISTNER, E. O., SHI, M. and WEINBERG, C. R. (2009). Using cases and parents to study multiplicative gene-by-environment interaction. *Am. J. Epidemiol.* **170** 393–400. https://doi.org/10.1093/aje/kwp118

KRAFT, P., YEN, Y. C., STRAM, D. O., MORRISON, J. and GAUDERMAN, W. J. (2007). Exploiting gene-environment interaction to detect genetic associations. *Hum. Hered.* **63** 111–119.

LUNDE, A., MELVE, K. K., GJESSING, H. K., SKJAERVEN, R. and IRGENS, L. M. (2007). Genetic and environmental influences on birth weight, birth length, head circumference, and gestational age by use of population-based parent–offspring data. *Am. J. Epidemiol.* **167** 734–741.

MALLIA, T., GRECH, A., HILI, A., CALLEJA-AGIUS, J. and PACE, N. P. (2017). Genetic determinants of low birth weight. *Minerva Ginecol.* **69** 631–643. https://doi.org/10.23736/S0026-4784.17.04050-3

PETRY, C. J., ONG, K. K. and DUNGER, D. B. (2007). Does the fetal genotype affect maternal physiology during pregnancy. *Trends Mol. Med.* **13** 414–421.

PIEGORSCH, W. W., WEINBERG, C. R. and TAYLOR, J. A. (1994). Non-hierarchical logistic models and case-only designs for assessing susceptibility in population-based case-control studies. *Stat. Med.* **13** 153–162.

PRENTICE, R. L. and PYKE, R. (1979). Logistic disease incidence models and case-control studies. *Biometrika* **66** 403–411. MR0556730 https://doi.org/10.1093/biomet/66.3.403

RADHA, V., KANTHIMATHI, S., ANJANA, R. M. and MOHAN, V. (2016). Genetics of gestational diabetes mellitus. *J. Pak. Med. Assoc.* **9 Suppl 1** S11–14.

SAFTLAS, A. F., BEYDOUN, H. and TRICHE, E. (2005). Immunogenetic determinants of preeclampsia and related pregnancy disorders: A systematic review. *Obstet. Gynecol.* **106** 162–172.

SHI, M., UMBACH, D. M. and WEINBERG, C. R. (2011). Family-based gene-by-environment interaction studies: Revelations and remedies. *Epidemiology* **22** 400–407.

SHI, M., UMBACH, D. M., VERMEULEN, S. H. and WEINBERG, C. R. (2008). Making the most of case-mother/control-mother studies. *Am. J. Epidemiol.* **168** 541–547.

SKARE, Ø., JUGESSUR, A., LIE, R. T., WILCOX, A. J., MURRAY, J. C., LUNDE, A., NGUYEN, T. T. and GJESSING, H. K. (2012). Application of a novel hybrid study design to explore gene-environment interactions in orofacial clefts. *Ann. Hum. Genet.* **76** 221–236.

UMBACH, D. and WEINBERG, C. (1997). Designing and analysing case-control studies to exploit independence of genotype and exposure. *Stat. Med.* **16** 1731–1743.

UMBACH, D. M. and WEINBERG, C. R. (2000). The use of case-parent triads to study joint effects of genotype and exposure. *Am. J. Hum. Genet.* **66** 251–261.

WANGLER, M. F., CHANG, A. S., MOLEY, K. H., FEINBERG, A. P. and DEBAUN, M. R. (2005). Factors associated with preterm delivery in mothers of children with Beckwith–Wiedemann syndrome: A case cohort study from the BWS registry. *Am. J. Med. Genet., Part A* **134** 187–191.

WEINBERG, C. R. and UMBACH, D. M. (2005). A hybrid design for studying genetic influences on risk of diseases with onset early in life. *Am. J. Hum. Genet.* **77** 627–636.

WEINBERG, C. R., WILCOX, A. J. and LIE, R. T. (1998). A log-linear approach to case-parent-triad data: Assessing effects of disease genes that act either directly or through maternal effects and that may be subject to parental imprinting. *Am. J. Hum. Genet.* **62** 969–978. https://doi.org/10.1086/301802

WILCOX, A. J., WEINBERG, C. R. and LIE, R. T. (1998). Distinguishing the effects of maternal and offspring genes through studies of "case-parent triads". *Am. J. Epidemiol.* **148** 893–901. https://doi.org/10.1093/oxfordjournals.aje.a009715

YANG, J. and LIN, S. (2013). Robust partial likelihood approach for detecting imprinting and maternal effects using case-control families. *Ann. Appl. Stat.* **7** 249–268. MR3086418 https://doi.org/10.1214/12-AOAS577

ZHANG, H., CHATTERJEE, N., RADER, D. and CHEN, J. (2018). Adjustment of nonconfounding covariates in case-control genetic association studies. *Ann. Appl. Stat.* **12** 200–221. MR3773391 https://doi.org/10.1214/17-AOAS1065

ZHANG, H., MUKHERJEE, B., ARTHUR, V., HU, G., HOCHNER, H. and CHEN, J. (2020). Supplement to "An efficient and computationally robust statistical method for analyzing case-control mother–offspring pair genetic association studies." https://doi.org/10.1214/19-AOAS1298SUPP.