Supplementary Material for"Imprinting and Maternal Effect Detection Using Partial Likelihood Based on Discordant Sibpair Data"

Fangyuan Zhang, Abbas Khalili, and Shili Lin

S1. Detailed Derivation of Probability for a DSP with Siblings

In the main text, the probability for a discordant sibpair with an arbitrary number of siblings is factored into three components (expressions (2)-(4) in main text). In the following, we provide the detailed derivation for the formula.

$$
\begin{aligned}
&P(M=m,F=f,C_1=c_1,C_2=c_2,C_i=c_i,D_i=d_i,i=3,\cdots \mid D_1=1,D_2=0)\\
&=P(M=m,F=f,C_1=c_1,C_2=c_2 \mid D_1=1,D_2=0)\\
&\times P(C_i=c_i,D_i=d_i,i=3,\cdots \mid M=m,F=f,C_1=c_1,C_2=c_2,D_1=1,D_2=0)\\
&=P(M=m,F=f,C_1=c_1,C_2=c_2 \mid D_1=1,D_2=0) \qquad (1)\\
&\times \prod_{i\geq 3} P(C_i=c_i|M=m,F=f)P(D_i=d_i|M=m,F=f,C_i=c_i). \qquad (2)
\end{aligned}
$$

The above expression holds because given parents' genotypes, different children's genotypes and disease status are independent. In particular, we note that expression (2) is the same as expression (3) in the main text. We then take a further look at expression (1).

$$
\begin{aligned}
&P(M=m,F=f,C_1=c_1,C_2=c_2 \mid D_1=1,D_2=0)\\
&=P(M=m,F=f,C_1=c_1 \mid D_1=1,D_2=0)P(C_2=c_2 \mid M=m,F=f,C_1=c_1,D_1=1,D_2=0)\\
&=P(M=m,F=f,C_1=c_1 \mid D_1=1,D_2=0)P(C_2=c_2 \mid M=m,F=f,D_2=0) \qquad (3)\\
&=P(M=m,F=f,C_1=c_1 \mid D_1=1,D_2=0)P(M=m,F=f,C_2=c_2 \mid D_1=1,D_2=0) \qquad (4)\\
&\times \frac{P(C_2=c_2 \mid M=m,F=f,D_2=0)}{P(M=m,F=f,C_2=c_2 \mid D_1=1,D_2=0)}. \qquad (5)
\end{aligned}
$$

Now note that (3) holds because of conditional independence again, and expresion (4) is the same as (2) in the main text. We then further check expression (5):

$$
\begin{aligned}
&\frac{P(C_2 = c_2 \mid M = m, F = f, D_2 = 0)}{P(M = m, F = f, C_2 = c_2 \mid D_1 = 1, D_2 = 0)} \\
&= \frac{P(C_2 = c_2 \mid M = m, F = f, D_2 = 0)P(D_1 = 1, D_2 = 0)}{P(M = m, F = f, C_2 = c_2, D_1 = 1, D_2 = 0)} \\
&= \frac{P(C_2 = c_2 \mid M = m, F = f, D_2 = 0)P(D_1 = 1, D_2 = 0)}{P(M = m, F = f, D_2 = 0)P(C_2 = c_2 \mid M = m, F = f, D_2 = 0)P(D_1 = 1 \mid C_2 = c_2, M = m, F = f, D_2 = 0)} \\
&= \frac{P(D_1 = 1, D_2 = 0)}{P(M = m, F = f, D_2 = 0)P(D_1 = 1 \mid M = m, F = f)} \qquad (6) \\
&= \frac{P(D_1 = 1, D_2 = 0)}{P(M = m, F = f)P(D_2 = 0 \mid M = m, F = f)P(D_1 = 1 \mid M = m, F = f)}. \qquad (7)
\end{aligned}
$$

Again, expression (6) holds because of conditional independence, and expression (7) is the same as (4) in the main text (4). Therefore, the probability of interest is factored into the products of expressions (4), (2) and (7), which correspond to expressions (2), (3), and (4), respectively, in the main text, completing the deviation of the probability.

S2. Calculation of Probabilities in Table 1.

Consider a candidate genetic marker with two alleles A and B, where A is the allele of interest, the variant allele, which may code for disease susceptibility or epigenetic effect. In a nuclear family, let F and M be the random variables denoting the number of A alleles carried by father and mother respectively, which can take values 0, 1, or 2, corresponding to genotype BB, AB or AA, respectively. Similarly, let C_i be the random variable denoting the number of A alleles, that is, the genotype of child i, $i = 1, 2, \cdots$. Specifically, C_1 and C_2 are designated for the affected and unaffected probands, respectively, through which the family is recruited, whereas $C_i, i = 3, \cdots$, are for the additional siblings, if any. $D_i,\ i = 1, 2, \cdots$, denote disease status of children (1 - affected; 0 - normal). Thus, $D_1 = 1$ and $D_2 = 0$.

In table 1, the formulas to calculate the joint probabilities are as follows:

$$
\begin{aligned}
&P(M = m, F = f, C_1 = c, D_1 = 1, D_2 = 0) \\
&= P(M = m, F = f)P(C_1 = c|M = m, F = f) \\
&\times P(D_1 = 1|M = m, F = f, C_1 = c)P(D_2 = 0|M = m, F = f), and
\end{aligned}
$$

$$
\begin{aligned}
&P(M = m, F = f, C_2 = c, D_1 = 1, D_2 = 0) \\
&= P(M = m, F = f)P(C_2 = c|M = m, F = f) \\
&\times P(D_2 = 0|M = m, F = f, C_2 = c)P(D_1 = 1|M = m, F = f).
\end{aligned}
$$

For all types other than type 8 (Table 1), if a child has one copy of the variant allele, the parental origin can be unambiguously identified, and hence the joint probability can be

easily obtained by extracting the relevant factors from the relative risk model for disease prevalence.

$$P(D=1|M=m,F=f,C=c)=\delta r_1^{I(c=1)}r_2^{I(c=2)}r_{im}^{I(c=1_m)}s_1^{I(m=1)}s_2^{I(m=2)}, \tag{8}$$

where the parameters: r_1 and r_2 denote the effect of one or two copies of an individual's own variant allele, r_{im} denotes imprinting effect, s_1 and s_2 denote the effect of one or two copies of the mother's variant allele, and δ is the phenocopy rate. The notation $c=1_m$ denotes that the child's genotype is heterozygous, where the variant allele is from mother. The indicator variable D denotes the disease status of a child (1 - affected; 0 - normal).We use μ_{mf}'s ($m=0,1,2,\ f=0,1,2$) to denote the mating type probabilities.

For example, in the familial genotype combination $(m,f,c)=(2,0,1)$,

$$\begin{aligned}
&P(M=2,F=0,C_1=1,D_1=1,D_2=0)\\
&=P(M=2,F=0)P(C_1=1|M=2,F=0)\\
&\times P(D_1=1|M=2,F=0,C_1=1)P(D_2=0|M=2,F=0)\\
&=\mu_{20}\delta r_1 s_2 r_{im}(1-\delta r_1 s_2 r_{im}),
\end{aligned}$$

and

$$\begin{aligned}
&P(M=2,F=0,C_2=1,D_1=1,D_2=0)\\
&=P(M=2,F=0)P(C_2=1|M=2,F=0)\\
&\times P(D_2=0|M=2,F=0,C_2=1)P(D_1=1|M=2,F=0)\\
&=\mu_{20}(1-\delta r_1 s_2 r_{im})\delta r_1 s_2 r_{im}.
\end{aligned}$$

For type 8, in which $(m,f,c)=(1,1,1)$, the variant allele carried by the child can be inherited either from the mother or the father with equal probabilities and, as such, the joint probability ends up being the summation of two probabilities weighted equally. We show the calculation of $P(M=1,F=1,C_1=1,D_1=1,D_2=0)$ as an example:

$$\begin{aligned}
&P(M=1,F=1,C_1=1,D_1=1,D_2=0)\\
&=P(M=1,F=1)P(C_1=1_m|M=1,F=1)\\
&\times P(D_1=1|M=1,F=1,C_1=1_m)P(D_2=0|M=1,F=1)\\
&+P(M=1,F=1)P(C_1=1_f|M=1,F=1)\\
&\times P(D_1=1|M=1,F=1,C_1=1_f)P(D_2=0|M=1,F=1)\\
&=1/4\mu_{11}\delta r_1 s_1(1+r_{im})1/4(4-\delta s_1-\delta r_1 s_1-\delta r_1 s_1 r_{im}-\delta r_2 s_1).
\end{aligned}$$

S3. Regularity Conditions and Proof of Theorem 1

The LIME$_{DSP}$ uses a multiplicative relative risk model for the disease prevalence are as given in (1) above. The vector of parameters of interest is denoted by

$$\boldsymbol{\theta}=(\delta,r_1,r_2,r_{im},s_1,s_2).$$

Let n^1_{mfc} and n^0_{mfc} denote the count of affected proband-parent triads and unaffected proband-parent triads with genotype $M = m$, $F = f$, and $C = c$, respectively. Similarly, let sn^1_{mfc} and sn^0_{mfc} denote the counts of affected additional sibling-parent triads and unaffected additional sibling-parent triads with genotype combination $M = m$, $F = f$ and $C = c$, respectively.

To make inference about $\boldsymbol{\theta}$, we use the partial log-likelihood

$$\begin{aligned} l_{par}(\boldsymbol{\theta}) \quad &= \sum_{m,f,c} \left\{ n^1_{mfc} \times \log[p_{mfc}(\boldsymbol{\theta})] + n^0_{mfc} \times \log[1 - p_{mfc}(\boldsymbol{\theta})] \right\} \\ &+ \sum_{m,f,c} \left\{ sn^1_{mfc} \times \log[q_{mfc}(\boldsymbol{\theta})] + sn^0_{mfc} \times \log[1 - q_{mfc}(\boldsymbol{\theta})] \right\} \\ &= l_{t1}(\boldsymbol{\theta}) + l_{t2}(\boldsymbol{\theta}). \end{aligned}$$

The effective total sample size, called n, in the partial log-likelihood $l_{par}(\boldsymbol{\theta})$, is computed as

$$\begin{aligned} n &= \sum_{m,f,c} [n^0_{mfc} + n^1_{mfc}] + \sum_{m,f,c} [sn^0_{mfc} + sn^1_{mfc}] \\ &= (N + N) + (sN^0_t + sN^1_t) \\ &= n_t + sn_t \end{aligned}$$

where N denotes the total number of independent families, and (sN^0_t, sN^1_t) are the total number of unaffected and affected siblings in all complete families, respectively. Hence n_t is the total number of probands children, and sn_t is the total number of additional siblings besides discordant sibpair.

The *maximum partial likelihood estimator* (MPLE) of $\boldsymbol{\theta}$ is denoted by

$$\widehat{\boldsymbol{\theta}}_n = \text{argmax}_{\boldsymbol{\theta}} \;\; l_{par}(\boldsymbol{\theta})$$

which is assumed to be obtained by solving the score-type equation

$$\frac{\partial l_{par}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = l'_{par}(\boldsymbol{\theta}) = l'_{t1}(\boldsymbol{\theta}) + l'_{t2}(\boldsymbol{\theta}) = \mathbf{0}.$$

We study the theoretical properties of $\widehat{\boldsymbol{\theta}}_n$, as the effective sample size $n = n_t + sn_t$ tends to infinity. We should note that here when $n \to \infty$, each of the sample sizes (n_t, sn_t) also tend to infinity, at the same rate, such that

$$\frac{n_t}{n} \longrightarrow 1 \;, \;\; \frac{sn_t}{n} \longrightarrow 1.$$

Clearly, this is under the assumption that both sums $\sum$ are present in the partial log-likelihood $l_{par}(\boldsymbol{\theta})$ defined above. If, however, there are no additional siblings, the theorem still holds and the proof is analogous.

Regularity Conditions

Let $\boldsymbol{\theta}_0$ be the true value of the parameter of interest. In what follows we denote

$$C_{r_n}(\boldsymbol{\theta}_0) = \{\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^6 : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leq r_n\}$$

as some neighborhood of $\boldsymbol{\theta}_0$, with radius r_n, where $r_n \to 0$, as n tends to infinity. Later on, we will see that this rate is $n^{-1/2}$. The regularity conditions are:

R1. The true value $\boldsymbol{\theta}_0$ of the parameter vector $\boldsymbol{\theta}$ is an interior point of the compact parameter space $\boldsymbol{\Theta}$.

R2. The cell probabilities $p_{mfc}(\boldsymbol{\theta})$ and $q_{mfc}(\boldsymbol{\theta})$ admit up to their third-order partial derivatives with respect to the elements of the parameter vector $\boldsymbol{\theta} = (\delta, r_1, r_2, r_{im}, s_1, s_2)$, for any $\boldsymbol{\theta} \in C_{r_n}(\boldsymbol{\theta}_0)$.

R3. The cell probabilities $p_{mfc}(\boldsymbol{\theta})$ and $q_{mfc}(\boldsymbol{\theta})$ are bounded away from the boundaries zero and one, at least for those $\boldsymbol{\theta} \in C_{r_n}(\boldsymbol{\theta}_0)$. Further, the partial derivatives of the cell probabilities, up to third order, are bounded by some constants, for any $\boldsymbol{\theta} \in C_{r_n}(\boldsymbol{\theta}_0)$.

R4. Identifiability: for any $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \boldsymbol{\Theta}$, $p_{mfc}(\boldsymbol{\theta}_1) = p_{mfc}(\boldsymbol{\theta}_2)$, $q_{mfc}(\boldsymbol{\theta}_1) = q_{mfc}(\boldsymbol{\theta}_2)$, for all (m, f, c) combinations, imply that $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$.

R5. The information matrix

$$I(\boldsymbol{\theta}) = -E\{l''_{par}(\theta)\} = -E\{\frac{\partial^2 l_{par}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}\}$$

is positive definite for any $\boldsymbol{\theta} \in C_{r_n}(\boldsymbol{\theta}_0)$.

We adopt the line of proof provided in Chanda (1954) and Lindsay (1980) to our partial likelihood context.

Proof of Theorem 1

Proof of Part (i) of Theorem 1. For simplicity in notation, we denote the vector of parameters of interest as $\boldsymbol{\theta} = (\delta, r_1, r_2, r_{im}, s_1, s_2) = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$. By the regularity Condition R2, for the first part of the partial log-likelihood, $l_{t1}(\boldsymbol{\theta})$, representing proband triads, we have that

$$\frac{\partial l_{t1}(\boldsymbol{\theta})}{\partial \theta_j} = l'_{t1,j}(\boldsymbol{\theta}) = l'_{t1,j}(\boldsymbol{\theta}_0) + \sum_{k=1}^{6} l''_{t1,jk}(\boldsymbol{\theta}_0)(\theta_k - \theta_k^0) + \frac{1}{2}\sum_{l,k}^{6} l'''_{t1,jkl}(\widetilde{\boldsymbol{\theta}})(\theta_k - \theta_k^0)(\theta_l - \theta_l^0) \quad (9)$$

for $j = 1, 2, \ldots, 6$, where $\widetilde{\boldsymbol{\theta}}$ is between $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta} \in C_{r_n}(\boldsymbol{\theta}_0)$; $l''_{t1,jk}(\cdot)$ and $l'''_{t1,jkl}(\cdot)$ are the second and third-order partial derivatives of the function $l_{t1}(\cdot)$, respectively. For $j, k, l = 1, 2, 3, 4, 5, 6$, we have

$$
\begin{aligned}
l'_{t1,j}(\boldsymbol{\theta}) &= \sum_{m,f,c} \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_j} \times \left\{ \frac{n^1_{mfc}}{p_{mfc}(\boldsymbol{\theta})} - \frac{n_{mfc} - n^1_{mfc}}{1 - p_{mfc}(\boldsymbol{\theta})} \right\} \\
l''_{t1,jk}(\boldsymbol{\theta}) &= \sum_{m,f,c} \frac{\partial^2 p_{mfc}(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_k} \times \left\{ \frac{n^1_{mfc}}{p_{mfc}(\boldsymbol{\theta})} - \frac{n_{mfc} - n^1_{mfc}}{1 - p_{mfc}(\boldsymbol{\theta})} \right\} \\
&- \sum_{m,f,c} \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_j} \times \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_k} \times \left\{ \frac{n^1_{mfc}}{[p_{mfc}(\boldsymbol{\theta})]^2} + \frac{n_{mfc} - n^1_{mfc}}{[1 - p_{mfc}(\boldsymbol{\theta})]^2} \right\} \\
l'''_{t1,jkl}(\boldsymbol{\theta}) &= \sum_{m,f,c} \frac{\partial^3 p_{mfc}(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_k \partial \theta_l} \times \left\{ \frac{n^1_{mfc}}{p_{mfc}(\boldsymbol{\theta})} - \frac{n_{mfc} - n^1_{mfc}}{1 - p_{mfc}(\boldsymbol{\theta})} \right\} \\
&- \sum_{m,f,c} \frac{\partial^2 p_{mfc}(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_k} \times \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_l} \times \left\{ \frac{n^1_{mfc}}{[p_{mfc}(\boldsymbol{\theta})]^2} + \frac{n_{mfc} - n^1_{mfc}}{[1 - p_{mfc}(\boldsymbol{\theta})]^2} \right\} \\
&- \sum_{m,f,c} \left[\frac{\partial^2 p_{mfc}(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_l} \times \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_k} + \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_j} \times \frac{\partial^2 p_{mfc}(\boldsymbol{\theta})}{\partial \theta_k \partial \theta_l} \right] \times \left\{ \frac{n^1_{mfc}}{[p_{mfc}(\boldsymbol{\theta})]^2} + \frac{n_{mfc} - n^1_{mfc}}{[1 - p_{mfc}(\boldsymbol{\theta})]^2} \right\} \\
&- \sum_{m,f,c} \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_j} \times \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_k} \times \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_l} \left\{ \frac{-2n^1_{mfc}}{[p_{mfc}(\boldsymbol{\theta})]^3} + \frac{2(n_{mfc} - n^1_{mfc})}{[1 - p_{mfc}(\boldsymbol{\theta})]^3} \right\}
\end{aligned}
$$

for any $\boldsymbol{\theta} \in C_{r_n}(\boldsymbol{\theta}_0)$.

For every triad type (m, f, c), denote the ratio

$$
r^1_{mfc} = \frac{n^1_{mfc}}{n_{mfc}}
$$

where $n_{mfc} = n^0_{mfc} + n^1_{mfc}$. The form of the partial log-likelihood $l_{par}(\boldsymbol{\theta})$ suggests that, for each triad type (m, f, c) and conditional on n_{mfc}, we have $n^1_{mfc} | n_{mfc} \sim Binomial(n_{mfc}, p_{mfc}(\boldsymbol{\theta}))$. By using a double conditional expectation technique, it is thus easy to see that $E(r^1_{mfc}) = p_{mfc}(\boldsymbol{\theta})$. Now, we have that

$$
\begin{aligned}
n^{-1} E\{l'_{t1,j}(\boldsymbol{\theta})\} &= 0 \\
-n^{-1} E\{l''_{t1,jk}(\boldsymbol{\theta})\} &= \sum_{m,f,c} \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_j} \times \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_k} \times \left\{ \frac{E(n_{mfc}/n)}{[p_{mfc}(\boldsymbol{\theta})][1 - p_{mfc}(\boldsymbol{\theta})]} \right\} = I_{t1,jk}(\boldsymbol{\theta})
\end{aligned}
$$

for any $\boldsymbol{\theta} \in C_{r_n}(\boldsymbol{\theta}_0)$, where $E(\cdot)$ is the expected value under the model with the parameter $\boldsymbol{\theta}$.

Further, by the regularity condition R3, for any $\boldsymbol{\theta} \in C_{r_n}(\boldsymbol{\theta}_0)$,

$$
\begin{aligned}
n^{-1} |l'''_{t1,jkl}(\boldsymbol{\theta})| &\leq \sum_{m,f,c} 2 \left| \frac{\partial^3 p_{mfc}(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_k \partial \theta_l} \right| + \sum_{m,f,c} \left| \frac{\partial^2 p_{mfc}(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_k} \times \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_l} \right| \times \left\{ \frac{(n_{mfc}/n)}{[p_{mfc}(\boldsymbol{\theta})][1 - p_{mfc}(\boldsymbol{\theta})]} \right\} \\
&+ \sum_{m,f,c} \left| \frac{\partial^2 p_{mfc}(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_l} \times \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_k} + \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_j} \times \frac{\partial^2 p_{mfc}(\boldsymbol{\theta})}{\partial \theta_k \partial \theta_l} \right| \times \left\{ \frac{(n_{mfc}/n)}{[p_{mfc}(\boldsymbol{\theta})][1 - p_{mfc}(\boldsymbol{\theta})]} \right\} \\
&+ 2 \sum_{m,f,c} \left| \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_j} \times \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_k} \times \frac{\partial p_{mfc}(\boldsymbol{\theta})}{\partial \theta_l} \right| \left\{ \frac{(n_{mfc}/n)}{[p_{mfc}(\boldsymbol{\theta})]^2} + \frac{(n_{mfc}/n)}{[1 - p_{mfc}(\boldsymbol{\theta})]^2} \right\} \\
&= O_p(1),
\end{aligned}
$$

which implies that $l'''_{t1,jkl}(\boldsymbol{\theta}) = O_p(n)$, for any $\boldsymbol{\theta} \in C_{r_n}(\boldsymbol{\theta}_0)$.

On the other hand, by the law of large numbers, we have that

$$r^1_{mfc} = \frac{n^1_{mfc}}{n_{mfc}} \xrightarrow{w.p.o} p_{mfc}(\boldsymbol{\theta}_0) \quad , \quad \frac{n_{mfc}}{n} \xrightarrow{w.p.o} E\left(\frac{n_{mfc}}{n}\right) = B_{mfc} \tag{10}$$

for some constant $0 < B_{mfc} < 1$, as $n \to \infty$, where w.p.o stands for with probability tending to one. Thus, using (10), as $n \to \infty$, we have

$$l'_{t1,j}(\boldsymbol{\theta}_0)/n \xrightarrow{w.p.o} 0 \ , \ l''_{t1,jk}(\boldsymbol{\theta}_0)/n \xrightarrow{w.p.o} I_{t1,jk}(\boldsymbol{\theta}_0) \ , \ l'''_{t1,jkl}(\boldsymbol{\theta}_0)/n = O_p(1). \tag{11}$$

for $j, k, l = 1, 2, \ldots, 6$.

By similar argument,s and under the regularity conditions R1-R5, for the remaining three terms of the partial log-likelihood, we have that

$$\begin{aligned}
n^{-1}E\{l'_{t2,j}(\boldsymbol{\theta})\} &= 0 \\
-n^{-1}E\{l''_{t2,jk}(\boldsymbol{\theta})\} &= \sum_{(m,f,c)} \frac{\partial q_{mfc}(\boldsymbol{\theta})}{\partial \theta_j} \times \frac{\partial q_{mfc}(\boldsymbol{\theta})}{\partial \theta_k} \times \left\{ \frac{E(sn_{mfc}/n)}{[q_{mfc}(\boldsymbol{\theta})][1 - q_{mfc}(\boldsymbol{\theta})]} \right\} = I_{t2,jk}(\boldsymbol{\theta}) \\
n^{-1}\{l'''_{t2,jkl}(\boldsymbol{\theta})\} &= O_p(1) \ \text{ as } \ n \to \infty.
\end{aligned}$$

Thus, similar to (11), as $n \to \infty$, we have that

$$l'_{t2,j}(\boldsymbol{\theta}_0)/n \xrightarrow{w.p.o} 0 \ , \ l''_{t2,jk}(\boldsymbol{\theta}_0)/n \xrightarrow{w.p.o} I_{t2jk}(\boldsymbol{\theta}_0) \ , \ l'''_{t2,jkl}(\boldsymbol{\theta}_0)/n = O_p(1),$$

for $j, k, l = 1, 2, \ldots, 6$.

Using the above results, we have that

$$l'_{par}(\boldsymbol{\theta}_0)/n \xrightarrow{w.p.o} 0 \ , \ l''_{par}(\boldsymbol{\theta}_0)/n \xrightarrow{w.p.o} \boldsymbol{I}(\boldsymbol{\theta}_0) \ , \ l'''_{par}(\boldsymbol{\theta}_0)/n = O_p(1) \tag{12}$$

as $n \to \infty$. Here $\boldsymbol{I}(\boldsymbol{\theta}_0)$ is a 6×6 information matrix constructed based on the $\{I_{t1,jk}(\boldsymbol{\theta}), I_{t2,jk}(\boldsymbol{\theta})\}$, for $j, k = 1, 2, \ldots, 6$.

Thus consider the score-type equation divided by the total sample size n, which leads to the equations

$$n^{-1}\sum_{k=1}^{6} l''_{par,jk}(\boldsymbol{\theta}_0)(\theta_k - \theta^0_k) = -n^{-1}l'_{par,j}(\boldsymbol{\theta}_0) - \frac{1}{2}n^{-1}\sum_{l,k=1}^{6} l'''_{par,jkl}(\widetilde{\boldsymbol{\theta}})(\theta_k - \theta^0_k)(\theta_l - \theta^0_l)$$

for $j = 1, \ldots, 6$. By expanding the summation on the left hand side and re-writing with respect to each $\theta_k - \theta^0_k$, we have that

$$\theta_k - \theta^0_k = \sum_{j=1}^{6}[\frac{-1}{n}l'_{par,j}(\boldsymbol{\theta}_0)] \times l^*_{par,jk}(\boldsymbol{\theta}_0) - \frac{1}{2}\sum_{l,r=1}^{6}\left[(\theta_r - \theta^0_r)(\theta_l - \theta^0_l)\left(\sum_{j=1}^{6}[\frac{1}{n}l'''_{par,jrl}(\widetilde{\boldsymbol{\theta}})] \times l^*_{par,jk}(\boldsymbol{\theta}_0)\right)\right] \tag{13}$$

for $k = 1, \ldots, 6$, where $l^*_{par,jk}(\boldsymbol{\theta}_0)$ are the elements of the inverse matrix $\Big(l''_{par,jk}(\boldsymbol{\theta}_0)/n; j, k = 1, \ldots, 6\Big)^{-1}$. By (12), the first term on the right hand side of the above equations tends to zero, as $n \to \infty$. This implies that the equations in (13) have at least one solution, in terms of $\theta_k - \theta^0_k$, that satisfies

$$\hat{\theta}_k - \theta^0_k \longrightarrow^p 0 \;\; ; \;\; k = 1, \ldots, 6,$$

as $n \to \infty$. Thus, there exists a solution, say, $\widehat{\boldsymbol{\theta}}_n$ of the score-type equation $l'_{par}(\boldsymbol{\theta}) = \mathbf{0}$ such that $\widehat{\boldsymbol{\theta}}_n \longrightarrow^p \boldsymbol{\theta}_0$, as $n \to \infty$.

Now we prove the uniqueness of such consistent estimator. Under the regularity conditions R1-R5, and consistency of $\hat{\boldsymbol{\theta}}_n$, we have that

$$\frac{1}{n} l''_{par}(\hat{\boldsymbol{\theta}}_n) + I(\boldsymbol{\theta}_0) = o_p(1) \tag{14}$$

as n tends to ∞, where $I(\boldsymbol{\theta}_0)$ is the positive definite information matrix. Let us assume that there exist two such consistent estimators, say, $\hat{\boldsymbol{\theta}}_{1n}$ and $\hat{\boldsymbol{\theta}}_{2n}$ of $\boldsymbol{\theta}_0$ that are the solutions of the score-type equation

$$l'_{par}(\boldsymbol{\theta}) = 0.$$

By the extension of Rolle's theorem to multivariate case, there exists a point $\tilde{\boldsymbol{\theta}}_n$ laying inside a hyper-cell with the vector $\hat{\boldsymbol{\theta}}_{1n} - \hat{\boldsymbol{\theta}}_{2n}$ as its diagonal, such that

$$l''_{par}(\tilde{\boldsymbol{\theta}}_n) = 0. \tag{15}$$

On the other hand, since $\hat{\boldsymbol{\theta}}_{1n}$ and $\hat{\boldsymbol{\theta}}_{2n}$ are consistent estimators, so is $\tilde{\boldsymbol{\theta}}_n$ and it must satisfy (14). But clearly (14) and (15) contradict. This implies that the consistent estimator $\hat{\boldsymbol{\theta}}_n$ is unique. This completes the proof of Part(i). ♠

The result of Lemma 1 below is used for proving Part (ii) of Theorem 1.

Lemma 1 *Under the regularity conditions R1-R5, we have that*

$$\frac{l'_{par}(\boldsymbol{\theta}_0)}{\sqrt{n}} \longrightarrow^d N(\mathbf{0}, \boldsymbol{I}(\boldsymbol{\theta}_0))$$

as $n \to \infty$.

Proof of Lemma 1. Consider the partial-score function

$$\begin{aligned}
\left.\frac{\partial l_{par}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = l'_{par}(\boldsymbol{\theta}_0) &= l'_{t1}(\boldsymbol{\theta}_0) + l'_{t2}(\boldsymbol{\theta}_0) \\
&= \sum_{m,f,c} \frac{n_{mfc} \times p'_{mfc}(\boldsymbol{\theta}_0)}{p_{mfc}(\boldsymbol{\theta}_0)[1 - p_{mfc}(\boldsymbol{\theta}_0)]} \times [r^1_{mfc} - p_{mfc}(\boldsymbol{\theta}_0)] \\
&+ \sum_{m,f,c} \frac{sn_{mfc} \times q'_{mfc}(\boldsymbol{\theta}_0)}{q_{mfc}(\boldsymbol{\theta}_0)[1 - q_{mfc}(\boldsymbol{\theta}_0)]} \times [s^1_{mfc} - q_{mfc}(\boldsymbol{\theta}_0)],
\end{aligned}$$

where $p'_{mfc}(\boldsymbol{\theta}_0)$ and $q'_{mfc}(\boldsymbol{\theta}_0)$ are the 6-dimensional vectors of the partial derivatives of the cell probabilities $p_{mfc}(\boldsymbol{\theta})$ and $q_{mfc}(\boldsymbol{\theta})$, with respect to $\boldsymbol{\theta}$, which are evaluated at the true $\boldsymbol{\theta}_0$. Also,

$$r^1_{mfc} = \frac{n^1_{mfc}}{n_{mfc}} \quad , \quad s^1_{mfc} = \frac{sn^1_{mfc}}{sn_{mfc}},$$

are the ratios of the number of cases among: proband (m, f, c) triads and additional (m, f, c) sibling triads respectively.

We first try to find the limiting distribution of $l'_{t1}(\boldsymbol{\theta}_0)/\sqrt{n}$, as $n \to \infty$. We have that

$$\frac{l'_{t1}(\boldsymbol{\theta}_0)}{\sqrt{n}} = \sum_{m,f,c} \frac{p'_{mfc}(\boldsymbol{\theta}_0)}{p_{mfc}(\boldsymbol{\theta}_0)[1 - p_{mfc}(\boldsymbol{\theta}_0)]} \times \sqrt{\frac{n_{mfc}}{n}} \times \sqrt{n_{mfc}}\ [r^1_{mfc} - p_{mfc}(\boldsymbol{\theta}_0)]$$

In what follows we use the Wald device. For any non-zero vector $\boldsymbol{v} \in \mathbb{R}^6$,

$$w_n(\boldsymbol{\theta}_0) = \frac{\boldsymbol{v}^\top l'_{t1}(\boldsymbol{\theta}_0)}{\sqrt{n}} = \sum_{m,f,c} \frac{u_{mfc}(\boldsymbol{\theta}_0)}{p_{mfc}(\boldsymbol{\theta}_0)[1 - p_{mfc}(\boldsymbol{\theta}_0)]} \times \sqrt{\frac{n_{mfc}}{n}} \times \sqrt{n_{mfc}}\ [r^1_{mfc} - p_{mfc}(\boldsymbol{\theta}_0)]$$

where $u_{mfc}(\boldsymbol{\theta}_0) = \boldsymbol{v}^\top p'_{mfc}(\boldsymbol{\theta}_0)$ is a scalar. Note that conditional on the n_{mfc}'s, the ratios r^1_{mfc}'s are independent, each having the conditional asymptotic distribution

$$\sqrt{n_{mfc}}\ [r^1_{mfc} - p_{mfc}(\boldsymbol{\theta}_0)] \longrightarrow^d N(0, p_{mfc}(\boldsymbol{\theta}_0)(1 - p_{mfc}(\boldsymbol{\theta}_0))$$

as $n \to \infty$. Note that since n_{mfc}'s are following a multinomial distribution, say, with the joint probability mass function $g(n_{mfc}; m, f, c)$, then

$$F_n(w) = P(w_n(\boldsymbol{\theta}_0) \le w) = \sum_{\{m,f,c:n_{mfc}=0\}}^{n_t} P(w_n(\boldsymbol{\theta}_0) \le w | n_{mfc}, m, f, c)\ g(n_{mfc}; m, f, c).$$

On the other hand, as $n \to \infty$, since $n_{mfc}/n \xrightarrow{p} E(n_{mfc}/n) = B_{mfc}$, for some constant $0 < B_{mfc} < 1$, then

$$(w_n(\boldsymbol{\theta}_0) | n_{mfc}, m, f, c) \longrightarrow^d N(0, \sigma^2(\boldsymbol{\theta}_0))$$

where

$$\sigma^2(\boldsymbol{\theta}_0) = \sum_{m,f,c} \frac{u^2_{mfc}(\boldsymbol{\theta}_0) \times B_{mfc}}{p_{mfc}(\boldsymbol{\theta}_0)(1 - p_{mfc}(\boldsymbol{\theta}_0))}.$$

Therefore, for $w \in \mathbb{R}$, as $n \to \infty$,

$$F_n(w) \longrightarrow \frac{1}{\sigma(\boldsymbol{\theta}_0)} \Phi\left(\frac{w}{\sigma(\boldsymbol{\theta}_0)}\right)$$

where $\Phi(\cdot)$ is the distribution function of the standard normal. This implies that

$$w_n(\boldsymbol{\theta}_0) \longrightarrow^d N(0, \sigma^2(\boldsymbol{\theta}_0))$$

as $n \to \infty$. Hence,

$$\frac{l'_{t1}(\boldsymbol{\theta}_0)}{\sqrt{n}} \longrightarrow^d N\left(\mathbf{0}, \sum_{m,f,c} \frac{[p'_{mfc}(\boldsymbol{\theta}_0)][p'_{mfc}(\boldsymbol{\theta}_0)]^\top \times B_{mfc}}{p_{mfc}(\boldsymbol{\theta}_0)(1-p_{mfc}(\boldsymbol{\theta}_0))}\right) \quad , \quad n \to \infty.$$

Similarly, we have

$$\frac{l'_{t2}(\boldsymbol{\theta}_0)}{\sqrt{n}} \longrightarrow^d N\left(\mathbf{0}, \sum_{m,f,c} \frac{[q'_{mfc}(\boldsymbol{\theta}_0)][q'_{mfc}(\boldsymbol{\theta}_0)]^\top \times C_{mfc}}{q_{mfc}(\boldsymbol{\theta}_0)(1-q_{mfc}(\boldsymbol{\theta}_0))}\right),$$

for some constants $0 < C_{mfc} < 1$, such that, as $n \to \infty$,

$$\frac{sn_{mfc}}{n} \longrightarrow^p C_{mfc}.$$

Thus, by the independence of the ratios r^1_{mfc} and s^1_{mfc}, as the effective sample size $n = n_t + sn_t$ tends to infinity, we have

$$\frac{l'_{par}(\boldsymbol{\theta}_0)}{\sqrt{n}} = \frac{l'_{t1}(\boldsymbol{\theta}_0)}{\sqrt{n}} + \frac{l'_{t2}(\boldsymbol{\theta}_0)}{\sqrt{n}} \longrightarrow^d N\left(\mathbf{0}, \boldsymbol{I}(\boldsymbol{\theta}_0)\right)$$

where $\boldsymbol{I}(\boldsymbol{\theta}_0) = \boldsymbol{I}_{t1}(\boldsymbol{\theta}_0) + \boldsymbol{I}_{t2}(\boldsymbol{\theta}_0)$, and

$$\begin{aligned}
\boldsymbol{I}_{t1}(\boldsymbol{\theta}_0) &= \sum_{m,f,c} \frac{[p'_{mfc}(\boldsymbol{\theta}_0)][p'_{mfc}(\boldsymbol{\theta}_0)]^\top \times B_{mfc}}{p_{mfc}(\boldsymbol{\theta}_0)(1-p_{mfc}(\boldsymbol{\theta}_0)}, \\
\boldsymbol{I}_{t2}(\boldsymbol{\theta}_0) &= \sum_{m,f,c} \frac{[q'_{mfc}(\boldsymbol{\theta}_0)][q'_{mfc}(\boldsymbol{\theta}_0)]^\top \times C_{mfc}}{q_{mfc}(\boldsymbol{\theta}_0)(1-q_{mfc}(\boldsymbol{\theta}_0)},
\end{aligned}$$

are 6×6-dimensional positive definite information matrices.

Hence, as $n \to \infty$, we have that

$$\frac{l'_{par}(\boldsymbol{\theta}_0)}{\sqrt{n}} \longrightarrow^d N(\mathbf{0}, \boldsymbol{I}(\boldsymbol{\theta}_0)). \tag{16}$$

This completes the proof of Lemma 1. ♠

Proof of Part (ii) of Theorem 1. Let $\widehat{\boldsymbol{\theta}}_n$ be the MPLE, which satisfies the score-type equation

$$l'_{par}(\widehat{\boldsymbol{\theta}}_n) = 0.$$

By the regularity conditions R1-R5, we have that

$$\begin{aligned}
\mathbf{0} &= \frac{1}{n}\, l'_{par}(\boldsymbol{\theta}_0) + \frac{1}{n}\, l''_{par}(\boldsymbol{\theta}_0)(1+o_p(1)) \times (\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \\
&= \frac{1}{n}\, l'_{par}(\boldsymbol{\theta}_0) + \left[\frac{1}{n}\, l''_{par}(\boldsymbol{\theta}_0) + \boldsymbol{I}(\boldsymbol{\theta}_0) - \boldsymbol{I}(\boldsymbol{\theta}_0)\right](1+o_p(1)) \times (\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)
\end{aligned}$$

where by (12) $l''_{par}(\boldsymbol{\theta}_0)/n + \boldsymbol{I}(\boldsymbol{\theta}_0) = o_p(1)$. Therefore, by the result of Lemma 1,

$$\sqrt{n}\,(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) = \boldsymbol{I}^{-1}(\boldsymbol{\theta}_0) \times \frac{l'_{par}(\boldsymbol{\theta}_0)}{\sqrt{n}} \longrightarrow^d N(0, \boldsymbol{I}^{-1}(\boldsymbol{\theta}_0)),$$

as $n \to \infty$. This completes the proof of Part(ii) of Theorem 1. ♠

S4. Estimation of Maternal Effect with the DSP Design without Additional Siblings

To analyze the information for detecting parent-of-origin effects, especially maternal effect, we take a closer look at p_{mfc} in the partial likelihood:

$$\begin{aligned} p_{mfc} &= \frac{P(D=1|m,f,c)P(D=0|m,f)}{P(D=1|m,f,c)P(D=0|m,f)+P(D=0|m,f,c)P(D=1|m,f)} \\ &= 1/\left(1+\frac{P(D=0|m,f,c)}{P(D=0|m,f)}/\frac{P(D=1|m,f,c)}{P(D=1|m,f)}\right). \end{aligned}$$

$$\begin{aligned} \frac{P(D=1|m,f,c)}{P(D=1|m,f)} &= \frac{\delta r_1^{I(C=1)} r_2^{I(C=2)} r_{im}^{I(C=1_m)} s_1^{I(M=1)} s_2^{I(M=2)}}{\sum_{c*} p(c*|m,f)\delta r_1^{I(C*=1)} r_2^{I(C*=2)} r_{im}^{I(C*=1_m)} s_1^{I(M=1)} s_2^{I(M=2)}} \\ &= \frac{r_1^{I(C=1)} r_2^{I(C=2)} r_{im}^{I(C=1_m)}}{\sum_{c*} p(c*|m,f) r_1^{I(C*=1)} r_2^{I(C*=2)} r_{im}^{I(C*=1_m)}}. \end{aligned} \tag{17}$$

$$\frac{P(D=0|m,f,c)}{P(D=0|m,f)} = \frac{1-\delta r_1^{I(c=1)} r_2^{I(c=2)} r_{im}^{I(c=1_m)} s_1^{I(m=1)} s_2^{I(m=2)}}{1-\sum_{c*} p(c*|m,f)\delta r_1^{I(c*=1)} r_2^{I(c*=2)} r_{im}^{I(c*=1_m)} s_1^{I(m=1)} s_2^{I(m=2)}}. \tag{18}$$

We can see that for maternal effect, (17) is totally independent of parameters s_1 and s_2. Though (18) includes maternal effect parameters, when there is only maternal effect, i.e. $r_1 = r_2 = r_{im} = 1$, maternal effect parameters will be canceled out again. Furthermore, when there are other effects besides maternal effect, only (F, M) belonging to $\{(1,2),(2,1),(1,0),(0,1),(1,1)\}$ is informative for (18), and if disease penetrance for these combinations with different offspring genotype are similar, for example, $P(D=1|M=1,F=2,C=1)$ is similar as $P(D=1|M=1,F=2,C=2)$, then the combination is again almost non-informative. On the other hand, most of child-parent genotype combinations are informative for detecting imprinting effect for both (17) and (18). This is consistent with the result from the simulation that the power to detect maternal effect is very low when only such discordant sibpairs without additional siblings are recruited, whereas when additional siblings are also recruited, the power will increase, as no term can be canceled.

S5. DSP design with missing father genotypes

In LIME proposed by Yang and Lin (2013), nuclear families with father's genotype missing can still contribute to the estimation of the parameters. However, as we elaborate in the following, LIME$_{DSP}$ cannot be generalized to the discordant sibpairs design with father's genotype missing. Following the same idea as in complete data, denote n_{mc}^1 as the count of affected proband-mother pairs with genotype $M = m$ and $C_1 = c$, and n_{mc}^0 as the

count of unaffected proband-mother pairs with genotype $M = m$ and $C_2 = c$. Let n_p denote the count of independent families. To keep it focused, we assume there are no additional siblings. Thus, the likelihood can be written as follows, where $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ denote the parameters of interest and the nuisance parameters, respectively. That is,

$$\begin{aligned} L(\boldsymbol{\theta}, \boldsymbol{\phi})_p &= \prod_{m,c} [p_{mc}^{n_{mc}^1} (1 - p_{mc})^{n_{mc}^0}] \prod_{m,c} S_{mc}^{n_{mc}^1 + n_{mc}^0} \\ &\times \prod_{j=1}^{n_p} \frac{P(M_j = m_j, C_{j1} = c_{j1}, C_{j2} = c_{j2})}{P(M_j = m_j, C_{j1} = c_{j1}) P(M_j = m_j, C_{j2} = c_{j2})} \frac{P(D_1 = 1, D_2 = 0)}{P(D_1 = 1 | m_j, c_{j2}) P(D_2 = 0 | m_j, c_{j1})}, \end{aligned} \tag{19}$$

where the j represents the j^{th} DSP in the data, and

$$p_{mc} = \frac{P(M = m, C_1 = c | D_1 = 1, D_2 = 0)}{P(M = m, C_1 = c | D_1 = 1, D_2 = 0) + P(M = m, C_2 = c | D_1 = 1, D_2 = 0)},$$

and the denominator is denoted as S_{mc}. However, we can rewrite the probability as

$$p_{mc} = \frac{1}{1 + \frac{P(M=m, C_1=c, D_1=1, D_2=0)}{P(M=m, C_2=c, D_1=1, D_2=0)}}.$$

Then, as we can see from Supplementary Table S7, p_{mc} still involves nuisance parameters, thus we cannot extract out a partial likelihood component to estimate parameters.

S6. Relative Efficiency of LIME$_{DSP}$ vs. LIME

To compare the relative efficiency of the LIME and LIME$_{DSP}$ study designs, we compare the "per individual" information when LIME$_{DSP}$ is applied to a D+2 design, with LIME to a T+3 study design, where a T+3 design refers to a case-parent/control-parent study design in which each family (either a case family or a control family) has 3 additional siblings. We chose to compare these two designs as the total number of individuals per family is equal to 6 in both designs. We vary the proportion of case families for the T+3 design from 0.025 to 0.975 by 0.025. Figures S18-25 are for disease model 1-8 under scenario 8, where the horizontal line is the information per individual for the D+2 design, while the circles represent that for the T+3 data. We can see that, as expected, a balanced setting, the proportion of case families being 0.5, is generally the most informative, in which case the D+2 design is not as efficient as the T+3 design. However, when such a balanced setting is not available, the D+2 design can be more efficient. This is especially true for making inference about association and imprinting effects. However, the T+3 design typically has more power than D+2 for inference about maternal effect, as we discussed earlier.

We further conducted a simulation study to illustrate empirically that LIME$_{DSP}$ can indeed be more powerful than LIME in settings in which there are very few control families. Specifically, for model 6 under scenario 8 (Table 2 in main text), we first applied LIME to 300 simulated T+3 families with the proportion of case families being 96.7% (i.e. 290 case families and 10 control families), and then applied LIME$_{DSP}$ to 300 simulated D+2 families. Note that both designs use the same number of families and each family contains

the same numbers of children and parents. In this case, LIME$_{DSP}$ achieves a higher power than LIME: 0.957 versus 0.856. This result is consistent with our theoretical calculation in this section. When balanced case-control family data are not available, using LIME$_{DSP}$ to analyze discordant sibpair data can be more efficient, and in fact necessary in the extreme situation when no control families are available at all.

References

Chanda, K. C. (1954), "A Note on the Consistency and Maxima of the Roots of Likelihood Equations", *Biometrika*, 41, 56–61.

Lindsay, B. G. (1980), "Nuisance Parameters, Mixture Models, and the Efficiency of Partial Likelihood Estimators", *Philosophical Transactions of the Royal Society of London A*, 296, 639–662.

Yang, J. and Lin, S. (2013), "Robust Partial Likelihood Approach for Detecting Imprinting and Maternal Effects Using Case-Control Families", *The Annals of Applied Statistics*, 7, 249–268.

SupplementaryTable S1. Top-20 SNPs having the smallest p-values for association with club foot using LIME_{DSP}

Rank	SNP	Chr	Position(BP)	Gene	-$\log_{10}$(P-value)
1	rs1023913	9	23003004	TOX3	4.7633
2	rs6040798	20	11602357		4.7631
3	rs1870488	6	63933078	WDR55	4.2773
4	rs292202	5	73582314	FAM53A	4.137
5	rs12523740	6	32897704		3.8777
6	rs10484209	4	37074039		3.8774
7	rs2953299	2	51852092		3.8746
8	rs1327992	6	4310124	CTB-32H22.1	3.7614
9	rs11594622	10	72580602		3.6976
10	rs17712426	10	83563646		3.6968
11	rs17035675	4	106457953		3.6754
12	rs6933121	6	79856243		3.6512
13	rs17141297	10	17580107		3.6244
14	rs12512863	4	24134430		3.6105
15	rs2650703	10	63236710	LOC101928781	3.5965
16	rs3115763	2	138763552	HNMT	3.5646
17	rs11980754	7	4408130		3.5394
18	rs1568717	15	61362446	RORA	3.5223
19	rs915895	6	32190216	KCND3	3.5093
20	rs2384549	12	115349867		4.9359

Supplementary Table S2. Top-20 SNPs having the smallest p-values for imprinting effect on club foot using LIME_{DSP}

Rank	SNP	Chr	Position(BP)	Gene	-$\log_{10}$(P-value)
1	rs1079295	5	5165951	MT1A	13.4218
2	rs2405941	18	73740843		13.2871
3	rs2320214	18	4420249	DLGAP1	12.4824
4	rs13384546	2	185616127	ZNF804A	12.2454
5	rs2145214	20	42237066	IFT52	11.9946
6	rs213134	17	32823258		11.7425
7	rs7162435	15	56121333	NEDD4	11.5518
8	rs6151826	5	80080680	MSH3	11.4768
9	rs2520121	16	26577301		11.4644
10	rs1224524	6	67250007		11.3491
11	rs10413941	19	49347707	PLEKHA4	11.1828
12	rs11610123	12	47500730	PCED1B	11.1069
13	rs11048527	12	26604100	ITPR2	11.1035
14	rs6785520	3	170991646	TNIK	10.9721
15	rs17117977	11	115130709		10.7654
16	rs13228877	7	34199973		10.6878
17	rs3743308	15	69563185	DRAIC	10.6850
18	rs11789529	9	130164412		10.5804
19	rs908296	2	9814639		10.4491
20	rs12223323	11	26298810	ANO3	10.3638

Supplementary Table S3. Top-20 SNPs having the smallest p-values for maternal effect on club foot using method LIME$_{DSP}$

Rank	SNP	Chr	Position(BP)	Gene	-log$_{10}$(P-value)
1	rs2384549	12	115349867		4.9359
2	rs3781503	10	121571506	INPP5F	4.9039
3	rs9446305	6	71598570	B3GAT2	4.5466
4	rs10224932	7	31035681		4.515
5	rs11766624	7	69887084	AUTS2	4.4982
6	rs585157	13	99045319	FARP1	4.467
7	rs9540648	13	34951551		4.3431
8	rs10499527	7	21243187		4.3245
9	rs1005391	4	16386448		4.2718
10	rs6711382	2	152531076	NEB	4.2556
11	rs7801891	7	17133513		4.2536
12	rs9818949	3	197683750	IQCG	4.2419
13	rs723636	6	160580493	SLC22A1	4.2334
14	rs2018193	1	153079071		4.215
15	rs10066164	5	13945188	DNAH5	4.2147
16	rs7546648	1	152931206		4.2143
17	rs17559561	4	132367852		4.1886
18	rs1529557	2	37898991		4.1799
19	rs12550249	8	13140608	DLC1	4.1429
20	rs17712426	10	83563646		3.6968

Supplementary Table S4. Top-20 SNPs having the smallest p-values for association with hypertension using LIME_{D+}

Rank	SNP	Chr	Position(BP)	Gene	-$\log_{10}$(P-value)
1	rs16892095	4	15518356	CC2D2A	15.65
2	rs11128437	3	75447270		15.48
3	rs4125931	4	49489497		15.35
4	rs2405219	18	731439945	SMIM21	15.26
5	rs2229188	7	92134309	CYP51A1	15.11
6	rs4702048	5	14750799	ANKH	14.44
7	rs12626631	21	45001813	HSF2BP	14.22
8	rs3734815	6	29694680	HLA-F	14.08
9	rs13202088	6	163174689	PACRG	13.64
10	rs52828135	15	unknown		13.50
11	rs6485742	11	12454075	PARVA	12.82
12	rs11843435	13	69479766		11.17
13	rs4707557	6	90362782	MDN1	11.16
14	rs7032988	9	91837409		9.93
15	rs2013347	17	22171189		8.73
16	rs11672918	19	8943393	ZNF558	8.62
17	rs13255458	8	41636070	ANK1	8.61
18	rs2272487	3	126733094	CHCHD6	8.41
19	rs2947658	3	125607009		8.07
20	rs12256916	10	38344894	ZNF33A	7.99

Supplementary Table S5. Top-20 SNPs having the smallest p-values for imprinting effect on hypertension using LIME_{D+}

Rank	SNP	Chr	Position(BP)	Gene	-$\log_{10}$(P-value)
1	rs16892095	4	15518356	CC2D2A	15.65
2	rs11128437	3	75447270		15.48
3	rs4125931	4	49489497		15.35
4	rs2405219	18	731439945	SMIM21	15.26
5	rs2229188	7	92134309	CYP51A1	15.11
6	rs4702048	5	14750799	ANKH	14.44
7	rs12626631	21	45001813	HSF2BP	14.22
8	rs3734815	6	29694680	HLA-F	14.08
9	rs13202088	6	163174689	PACRG	13.64
10	rs52828135	15	unknown		13.50
11	rs6485742	11	12454075	PARVA	12.82
12	rs11843435	13	69479766		11.17
13	rs4707557	6	90362782	MDN1	11.16
14	rs7032988	9	91837409		9.93
15	rs2013347	17	22171189		8.73
16	rs11672918	19	8943393	ZNF558	8.62
17	rs13255458	8	41636070	ANK1	8.61
18	rs2272487	3	126733094	CHCHD6	8.41
19	rs2947658	3	125607009		8.07
20	rs12256916	10	38344894	ZNF33A	7.99

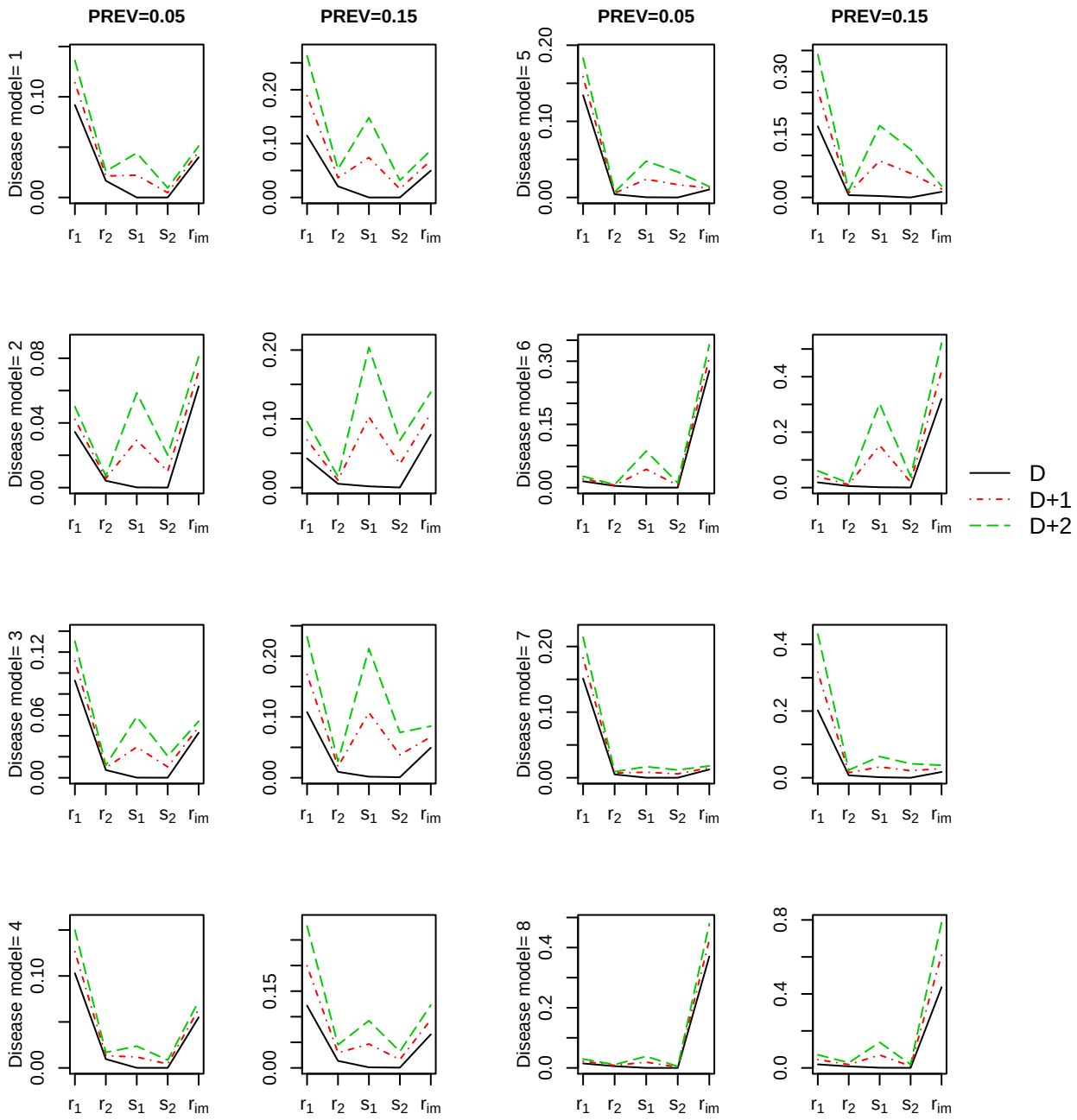
Supplementary Table S6. Top-20 SNPs having the smallest p-values for maternal effect on hypertension using LIME_{D+}

Rank	SNP	Chr	Position(BP)	Gene	-$\log_{10}$(P-value)
1	rs2272487	3	126451936	CHCHD6	8.44
2	rs9852584	3	126445456	CHCHD6	6.26
3	rs13230531	7	6114558	CHCHD6	5.52
4	rs17631957	14	81755544	STON2	5.49
5	rs820866	5	73978700		5.43
6	rs6086342	20	8096104		5.23
7	rs7741727	6	132069916	ENPP3	5.19
8	rs1370656	2	178607997	PDE11A	5.18
9	rs7133914	12	40702910	LRRK2	5.16
10	rs17601580	6	132061419	ENPP3	5.07
11	rs3856154	1	225565014	DNAH14	5.03
12	rs2165661	11	100142833	CNTN5	4.99
13	rs12368599	12	12908793	GPRC5A	4.92
14	rs17158657	15	84405464	ADAMTSL3	4.90
15	rs16832191	3	120944943	STXBP5L	4.88
16	rs3205144	3	172349215	NCEH1	4.82
17	rs4813864	20	8515840	PLCB1	4.78
18	rs17460330	4	36338943	DTHD1	4.76
19	rs10209069	2	153384254	FMNL2	4.71
20	rs390878	4	103213241	SLC39A8	4.67

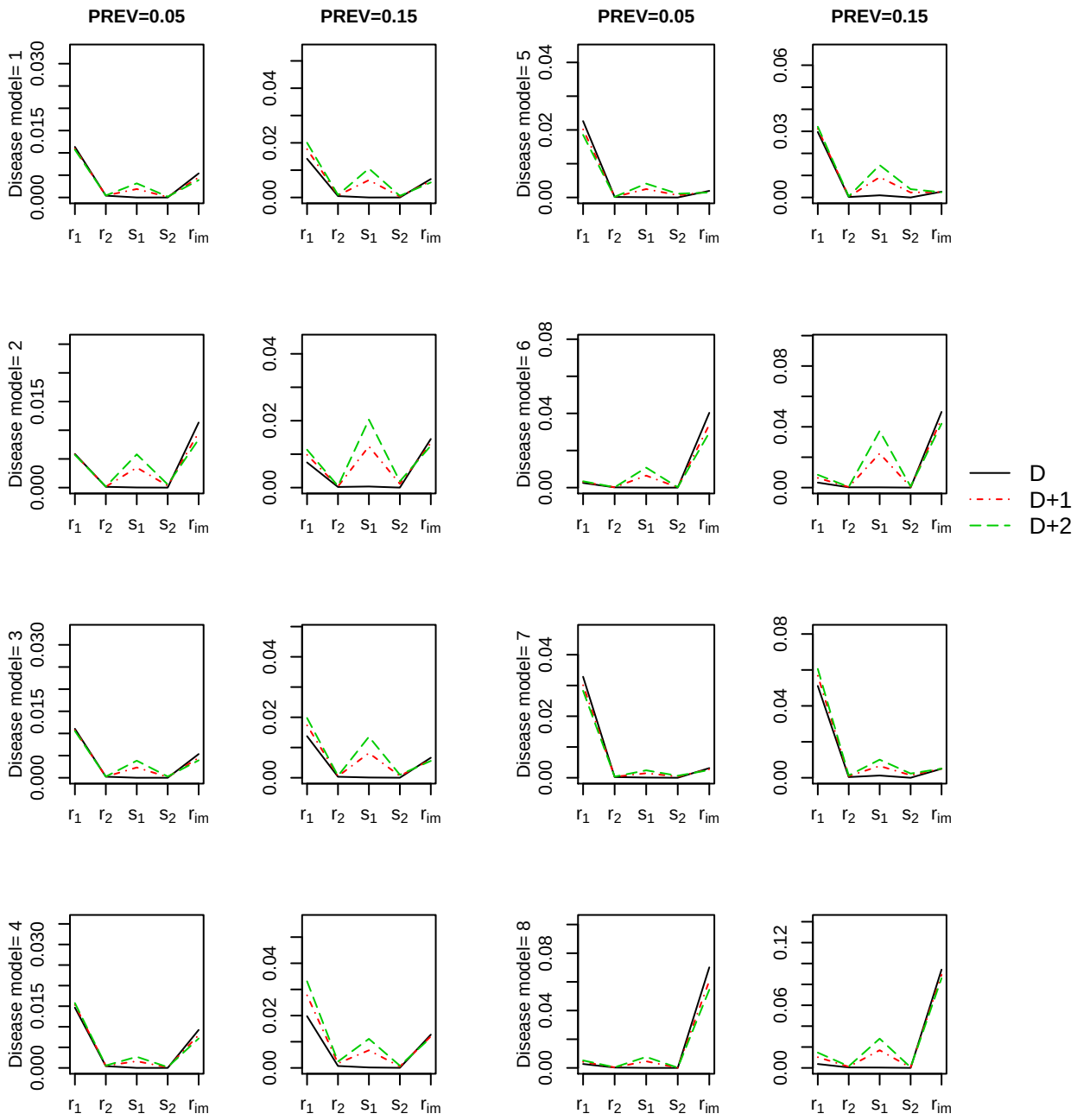
Supplementary Table S7. Joint probabilities of $P(M = m, C_1 = c, D_1 = 1, D_2 = 0)$ and $P(M = m, C_2 = c, D_1 = 1, D_2 = 0)$

Type	m	c	$P(M = m, C_1 = c, D_1 = 1, D_2 = 0)$
1	0	0	$\mu_{00}(1-\delta)\delta + \frac{1}{4}\mu_{01}\delta(2-\delta-\delta r_1)^a$
2	0	1	$\frac{1}{4}\mu_{01}\delta r_1(2-\delta r_1-\delta) + \mu_{02}(1-\delta r_1)\delta r_1$
3	1	0	$\frac{1}{4}\mu_{10}\delta s_1(2-\delta s_1-\delta s_1 r_1 r_{im})$
			$+\frac{1}{16}\mu_{11}\delta s_1(4-\delta s_1-\delta s_1 r_1(1+r_{im})-\delta s_1 r_2)$
4	1	1	$\frac{1}{4}\mu_{10}\delta s_1 r_1 r_{im}(2-\delta s_1-\delta s_1 r_1 r_{im})$
			$+\frac{1}{16}\mu_{11}\delta s_1 r_1(1+r_{im})(4-\delta s_1-\delta s_1 r_1-\delta s_1 r_1 r_i m-\delta r_2 s_1)$
			$+\frac{1}{4}\mu_{12}\delta r_1 s_1(2-\delta r_1 s_1-\delta r_2 s_1)$
5	1	2	$\frac{1}{16}\mu_{11}\delta s_1 r_2(4-\delta s_1-\delta s_1 r_1(1+r_{im})-\delta s_1 r_2)$
			$+\frac{1}{4}\mu_{12}\delta s_1 r_2(2-\delta s_1 r_1-\delta s_1 r_2)$
6	2	1	$\mu_{20}(1-\delta s_2 r_1 r_{im})\delta s_2 r_1 r_{im}$
			$+\frac{1}{4}\mu_{21}\delta s_2 r_1 r_{im}(2-\delta s_2 r_1 r_{im}-\delta s_2 r_2)$
7	2	2	$\frac{1}{4}\mu_{21}\delta s_2 r_2(2-\delta s_2 r_1 r_{im}-\delta s_2 r_2)+\mu_{22}(1-\delta s_2 r_2)\delta r_2 s_2$
Type	m	c	$P(M = m, C_2 = c, D_1 = 1, D_2 = 0)$
1	0	0	$\mu_{00}(1-\delta)\delta + \frac{1}{4}\mu_{01}(1-\delta)\delta(1+r_1)$
2	0	1	$\frac{1}{4}\mu_{01}(1-\delta r_1)\delta(1+r_1)+\mu_{02}(1-\delta r_1)\delta r_1$
3	1	0	$\frac{1}{4}\mu_{10}(1-\delta s_1)\delta s_1(1+r_1 r_{im})$
			$+\frac{1}{16}\mu_{11}(1-\delta s_1)\delta s_1(1+r_2+r_1(1+r_{im}))$
4	1	1	$\frac{1}{4}\mu_{10}(1-\delta s_1 r_1 r_{im})\delta s_1(1+r_1 r_{im})$
			$+\frac{1}{16}\mu_{11}[2-\delta r_1 s_1(1-r_{im})]\delta s_1(1+r_1(1+r_{im})+r_2)$
			$+\frac{1}{4}\mu_{12}(1-\delta r_1 s_1)\delta s_1(r_1+r_2)$
5	1	2	$\frac{1}{16}\mu_{11}(1-\delta s_1 r_2)\delta s_1(1+r_2+r_1(1+r_{im}))$
			$+\frac{1}{4}\mu_{12}(1-\delta s_1 r_2)\delta s_1(r_1+r_2)$
6	2	1	$\mu_{20}(1-\delta s_2 r_1 r_{im})\delta s_2 r_1 r_{im}$
			$+\frac{1}{4}\mu_{21}(1-\delta s_2 r_1 r_{im})\delta s_2(r_2+r_1 r_{im})$
7	2	2	$\frac{1}{4}\mu_{21}(1-\delta s_2 r_2)\delta s_2(r_1 r_{im}+r_2)+\mu_{22}(1-\delta s_2 r_2)\delta r_2 s_2$

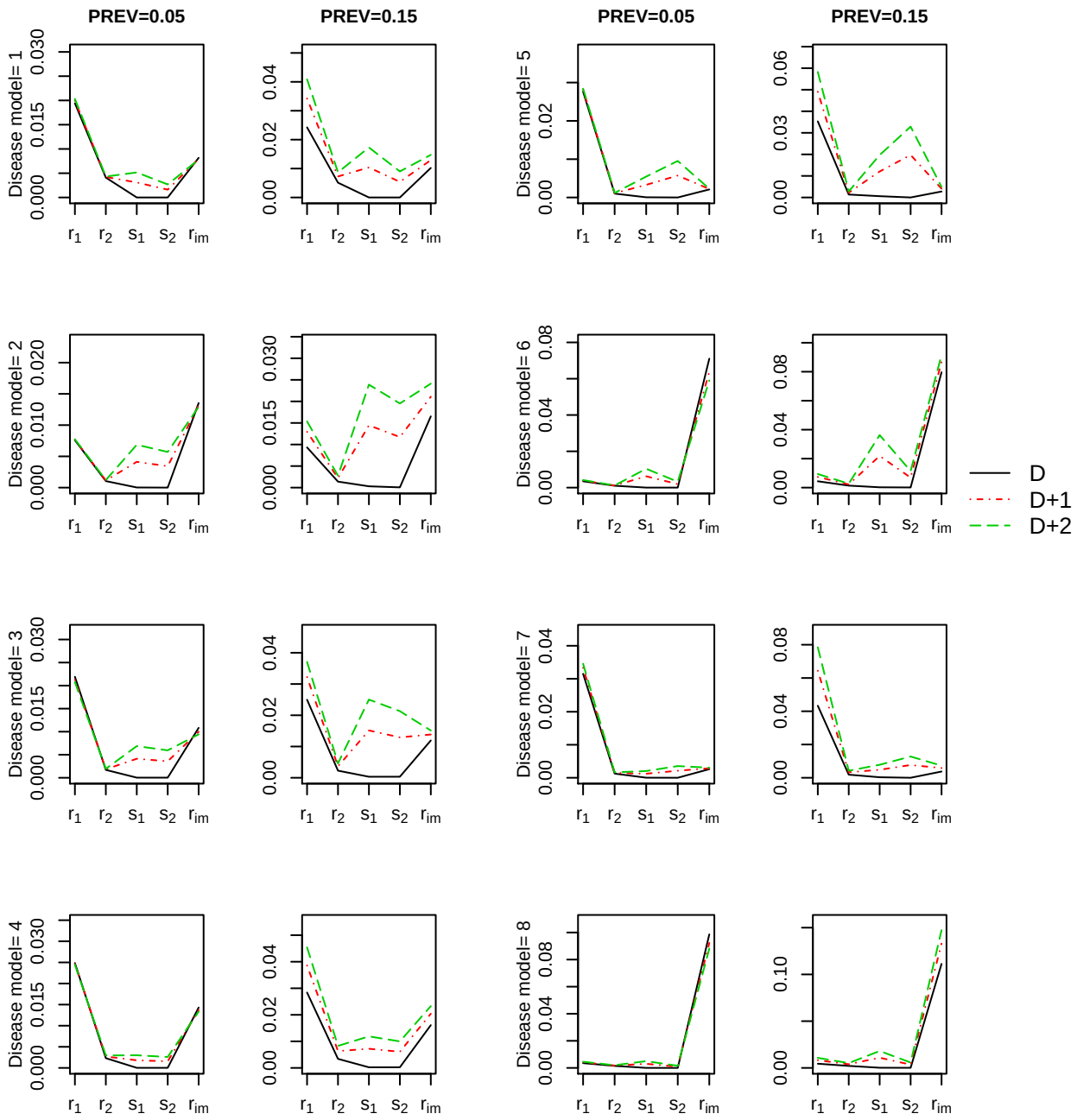
Note: $^a r_1$: relative risk of carrying one variant allele; r_2: relative risk of carry ing two variant alleles; r_{im}: imprinting effect parameter with a single variant allele from mother; s_1: maternal effect with mother carrying one variant allele; s_2: maternal effect with mother carrying two variant allele. In addition, mating type probability of $(M, F) = (m, f)$ is denoted by μ_{ij}.



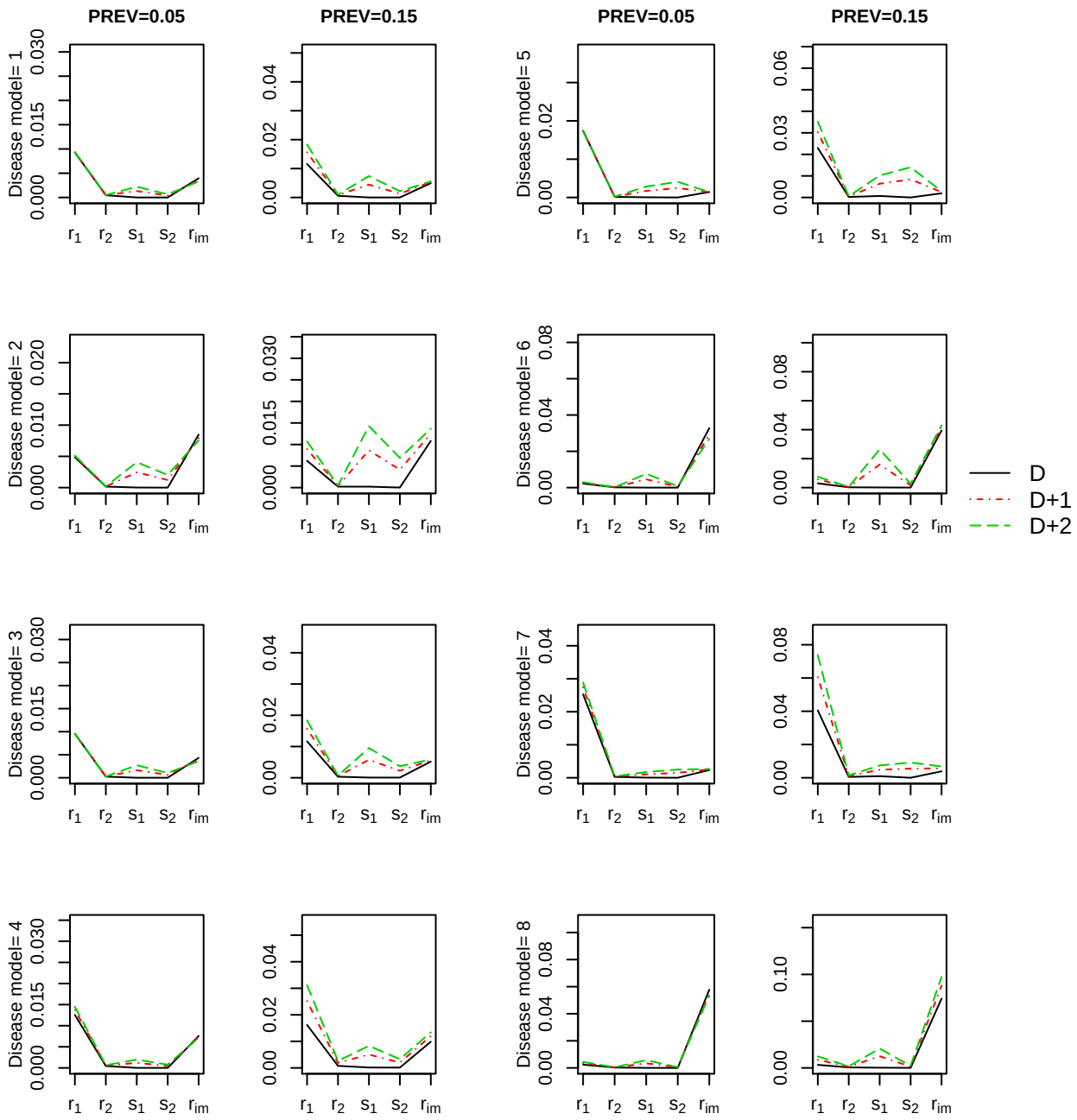
Supplementary Figure S1. Information content per family for 8 disease models and two PREVs when HWE holds and MAF is 0.3. Each curve provides the information for estimating one of the 5 parameters, for data types D, $D+1$ and $D+2$.



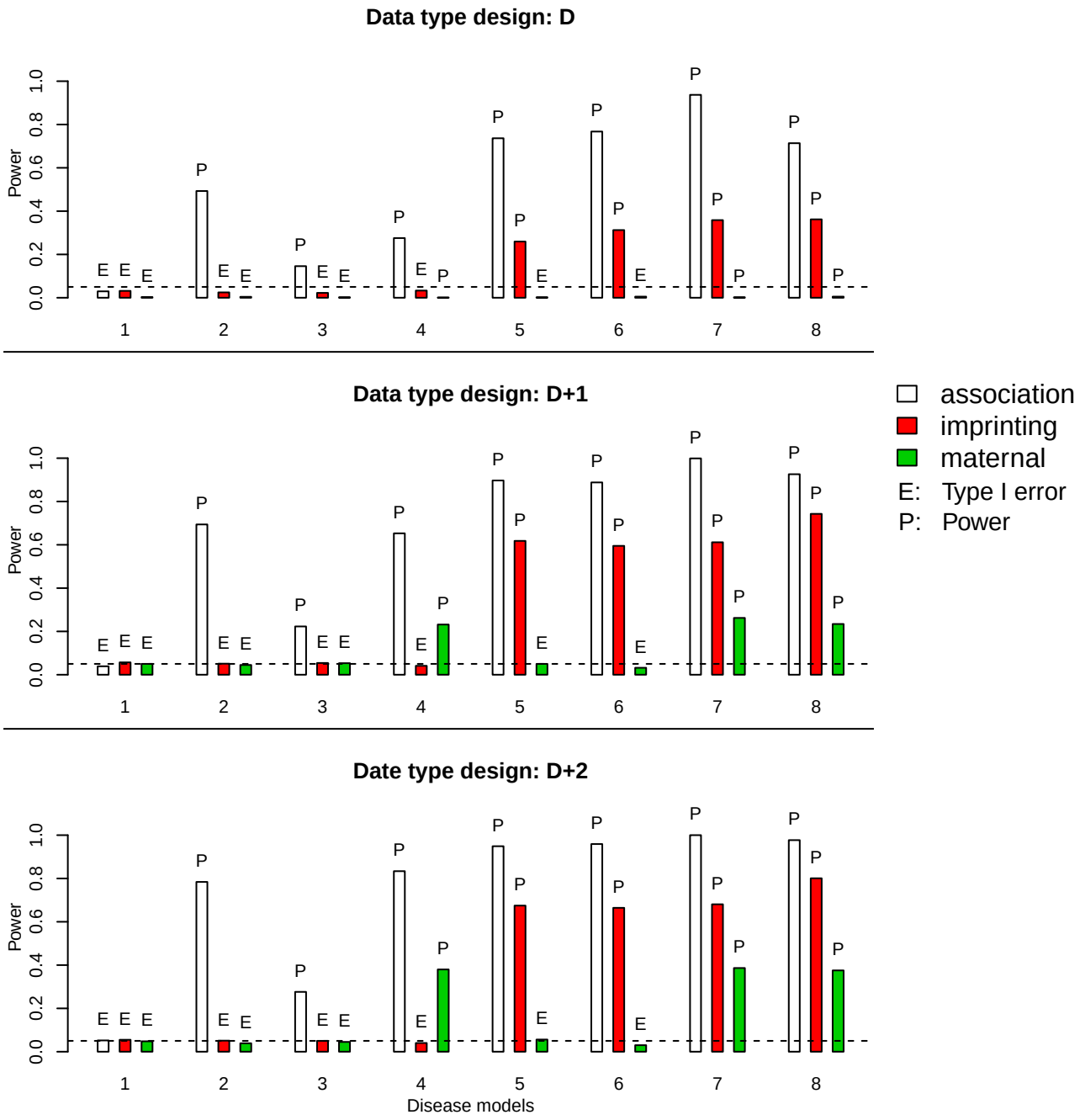
Supplementary Figure S2. Information content per individual for 8 disease models and two PREVs when HWE holds and MAF is 0.1. Each curve provides the information for estimating one of the 5 parameters, for data types D, $D+1$ and $D+2$.



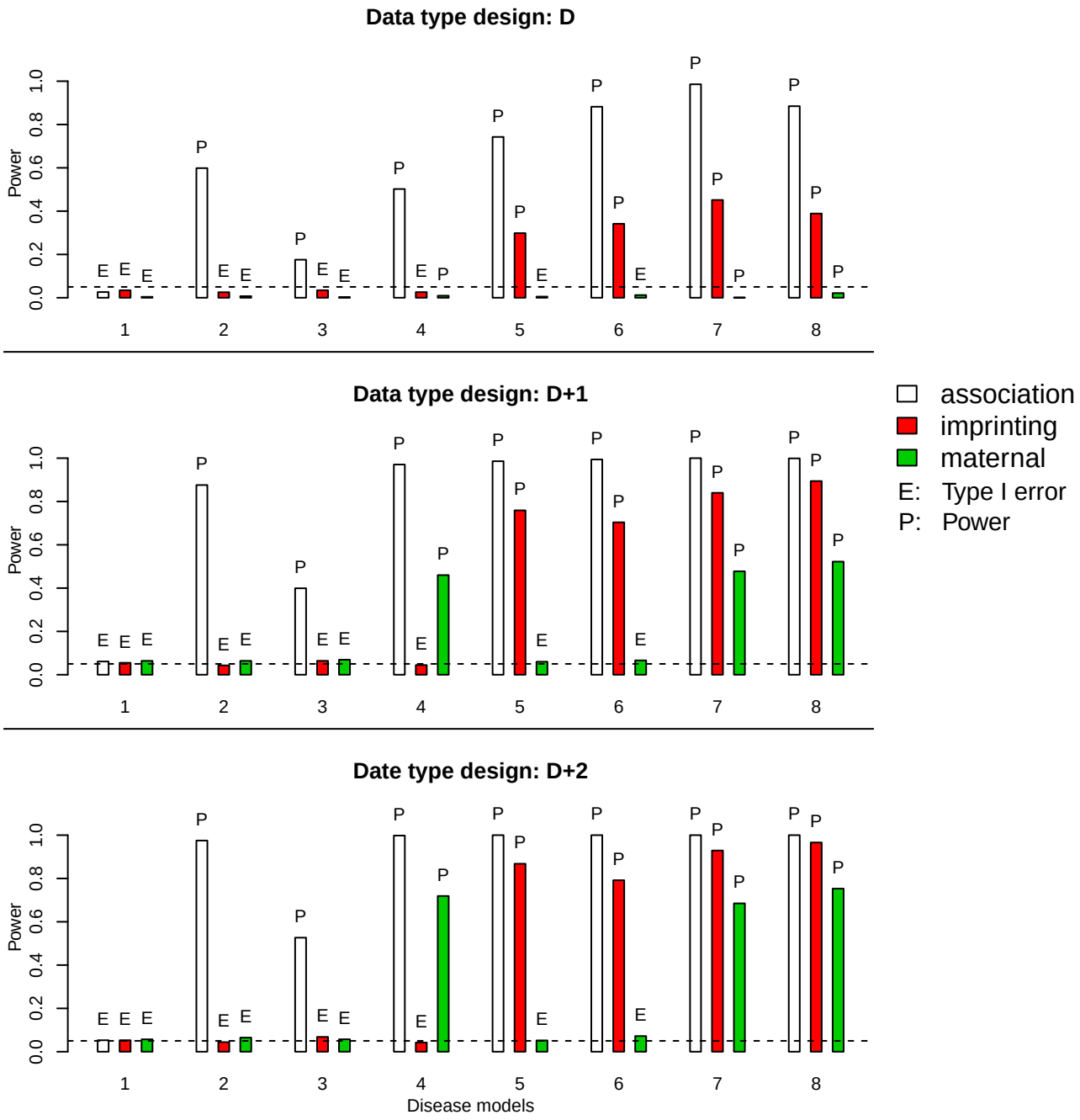
Supplementary Figure S3. Information content per individual for 8 disease models and two PREVs when HWE does not hold and MAF is 0.3. Each curve provides the information for estimating one of the 5 parameters, for data types D, $D+1$ and $D+2$.



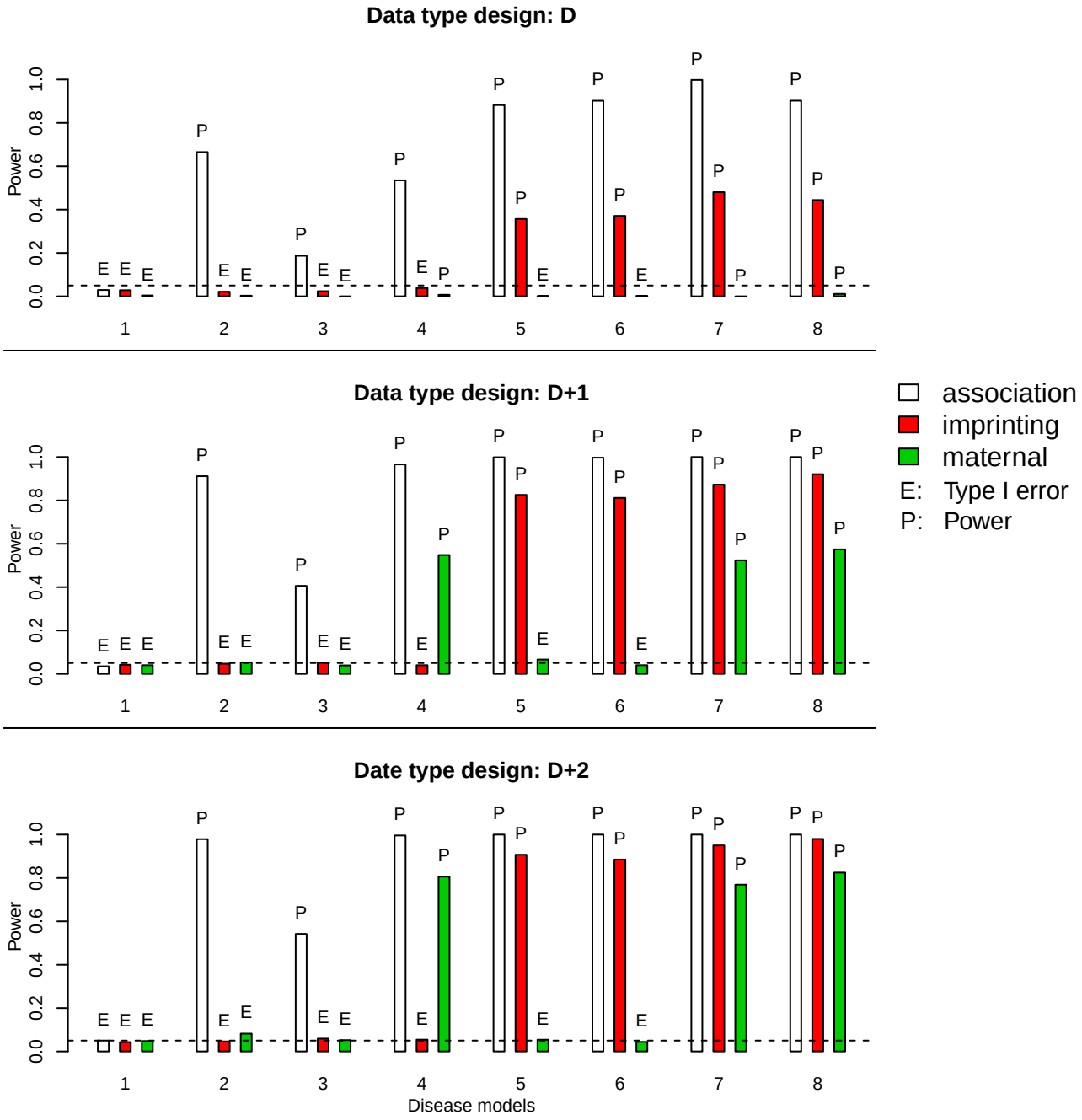
Supplementary Figure S4. Information content per individual for 8 disease models and two PREVs when HWE does not hold and MAF is 0.1. Each curve provides the information for estimating one of the 5 parameters, for data types D, $D+1$ and $D+2$.



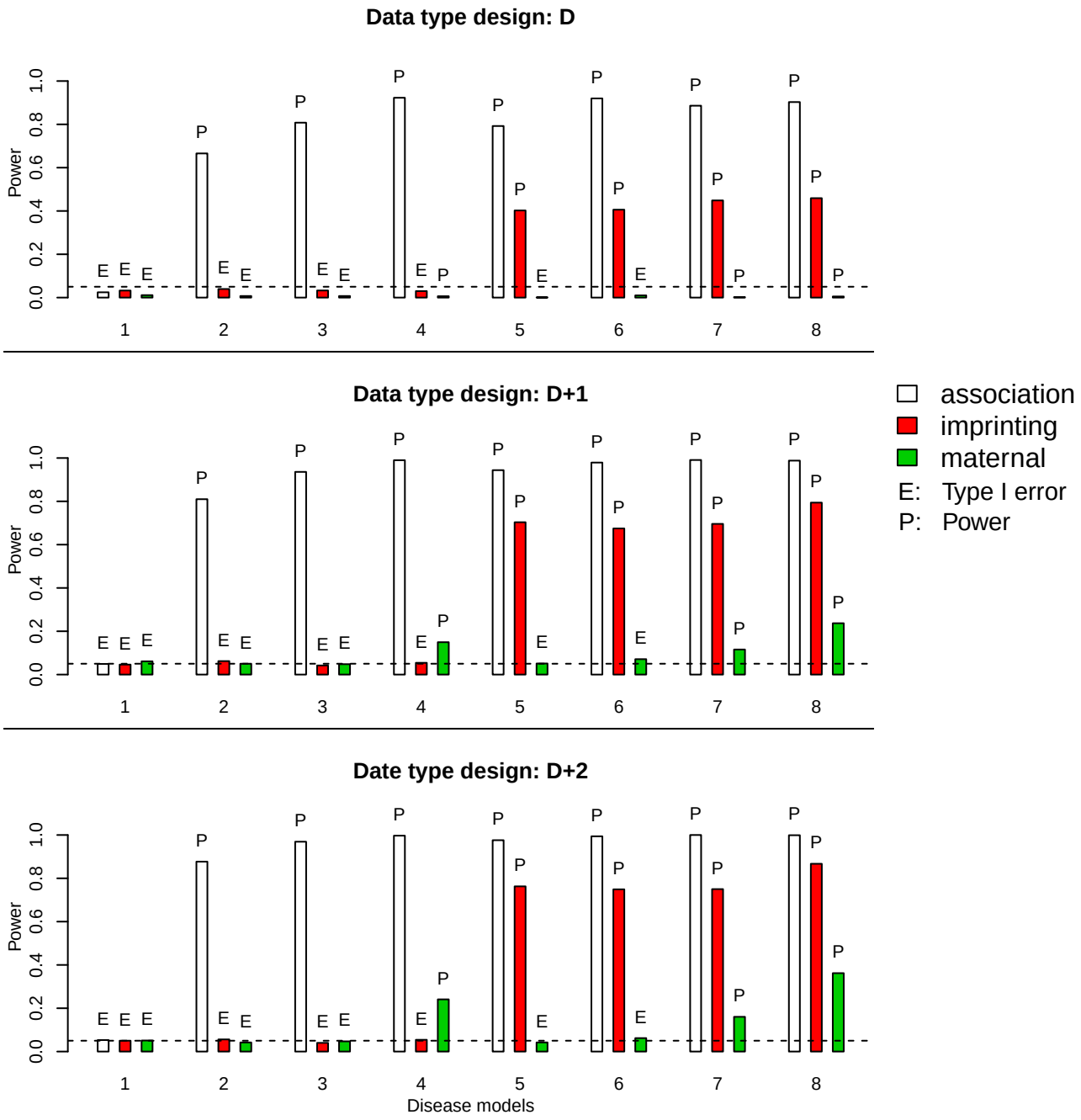
Supplementary Figure S5. Type I error rate and power of LIME$_{DSP}$ under 8 disease models and scenario 2 as given in Table 2. Three rows represent three data types: D, $D+1$ and $D+2$. The bars of color white, red and green refer to association, imprinting effect and maternal effect. The horizontal line marks the nominal a level of 0.05.



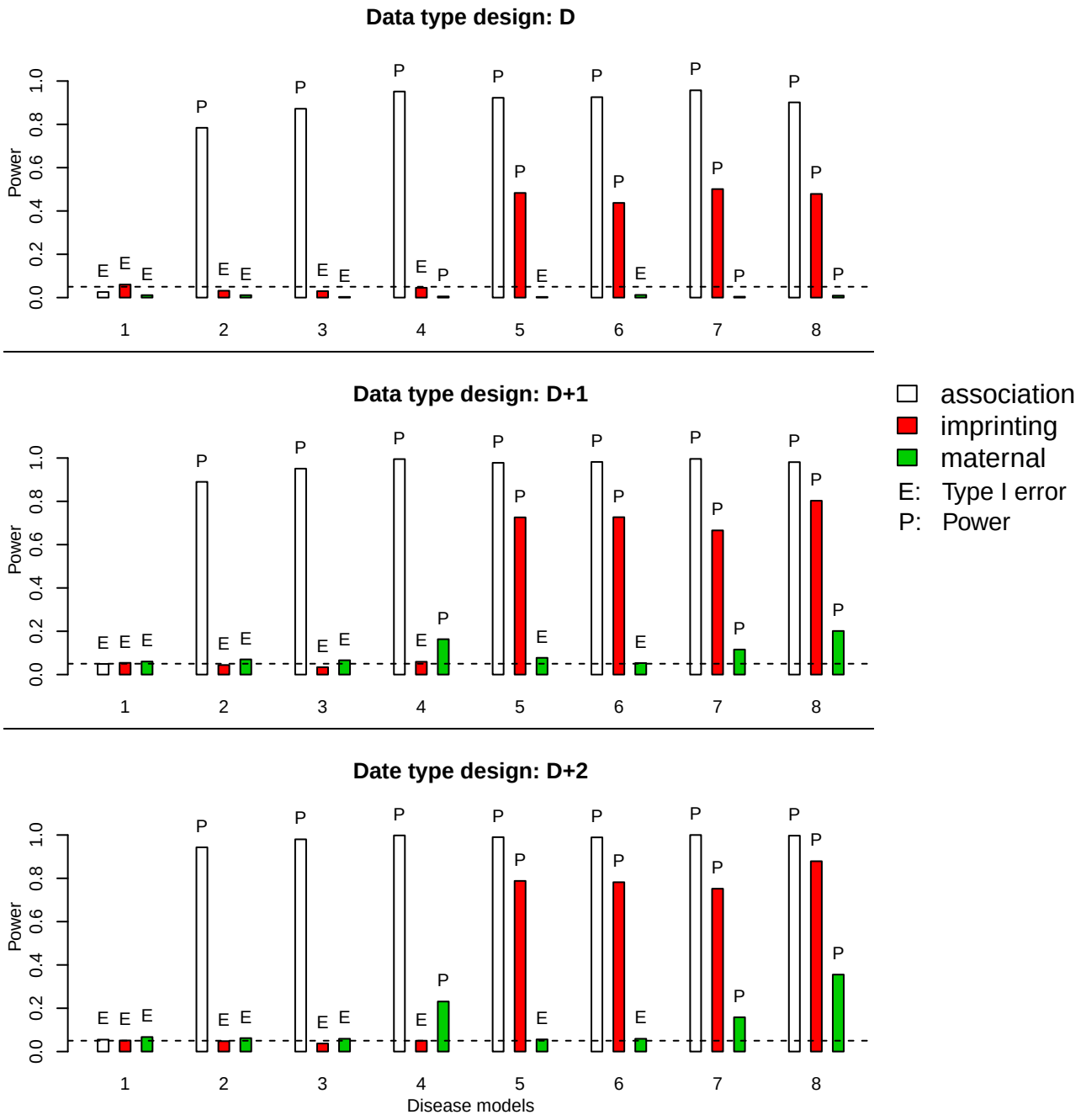
Supplementary Figure S6. Type I error rate and power of LIME$_{DSP}$ under 8 disease models and scenario 3 as given in Table 2. Three rows represent three data types: D, $D+1$ and $D+2$. The bars of color white, red and green refer to association, imprinting effect and maternal effect. The horizontal line marks the nominal a level of 0.05.



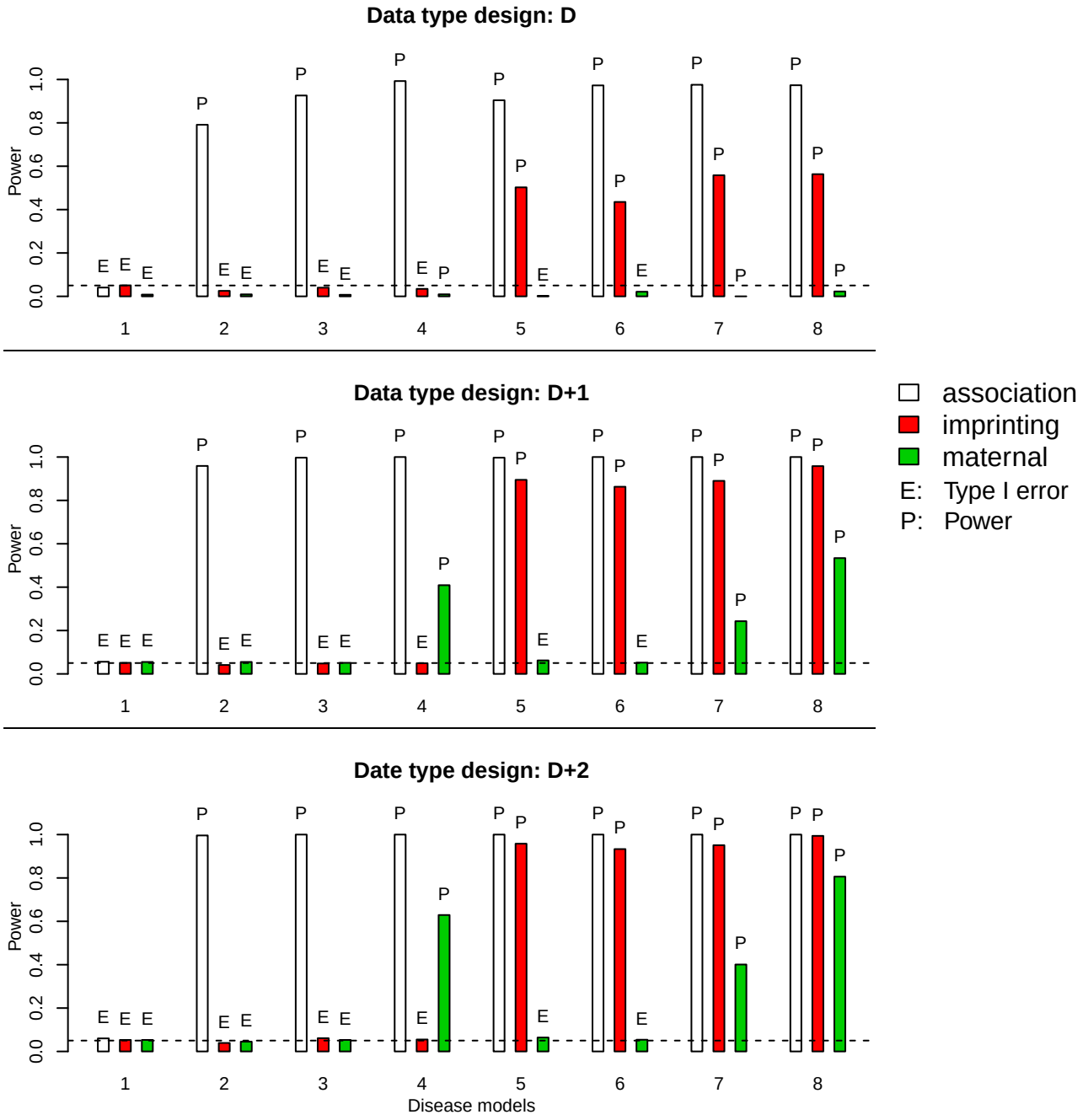
Supplementary Figure S7. Type I error rate and power of LIME$_{DSP}$ under 8 disease models and scenario 4 as given in Table 2. Three rows represent three data types: D, $D+1$ and $D+2$. The bars of color white, red and green refer to association, imprinting effect and maternal effect. The horizontal line marks the nominal a level of 0.05.



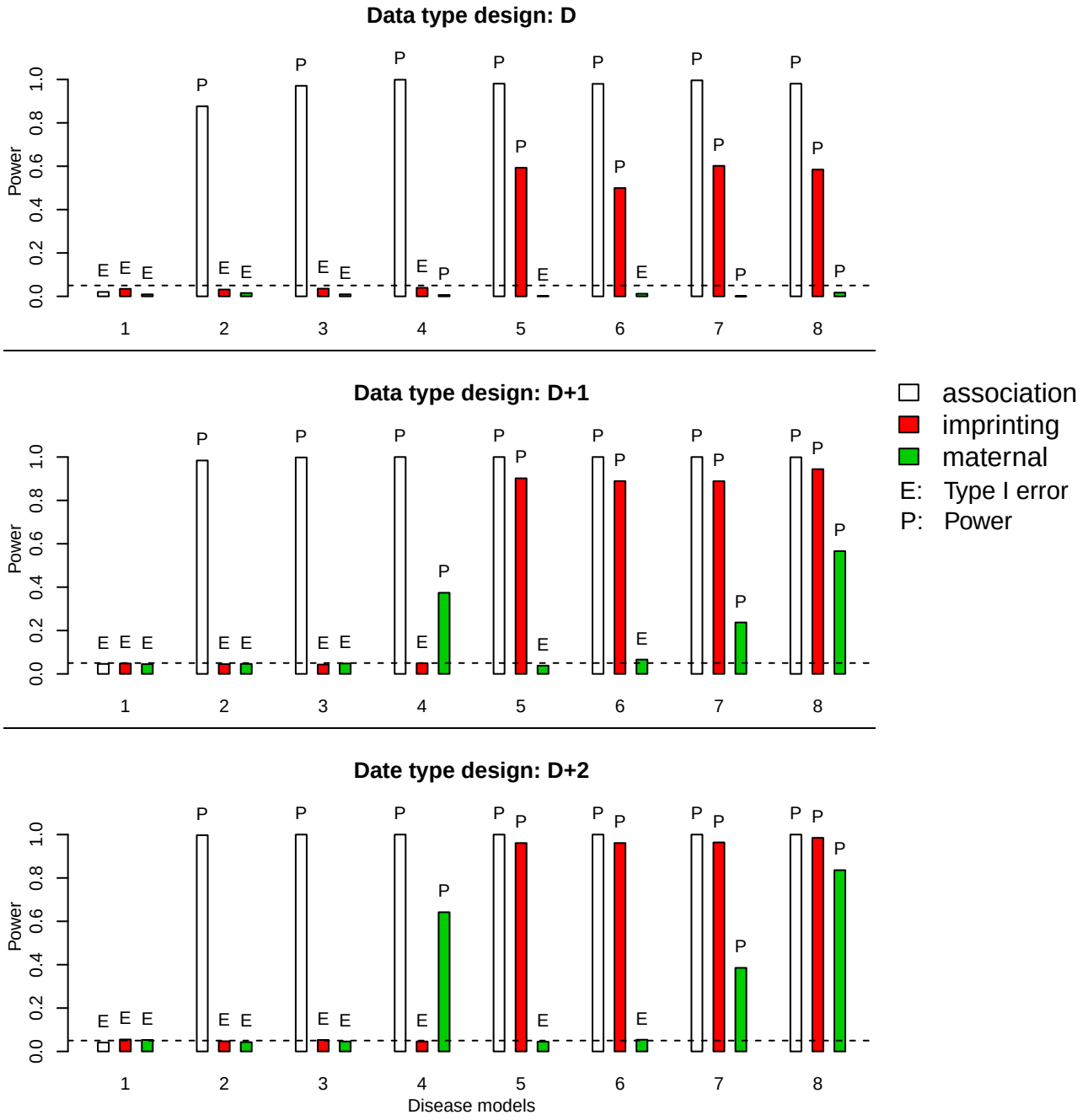
Supplementary Figure S8. Type I error rate and power of LIME$_{DSP}$ under 8 disease models and scenario 5 as given in Table 2. Three rows represent three data types: D, $D+1$ and $D+2$. The bars of color white, red and green refer to association, imprinting effect and maternal effect. The horizontal line marks the nominal a level of 0.05.



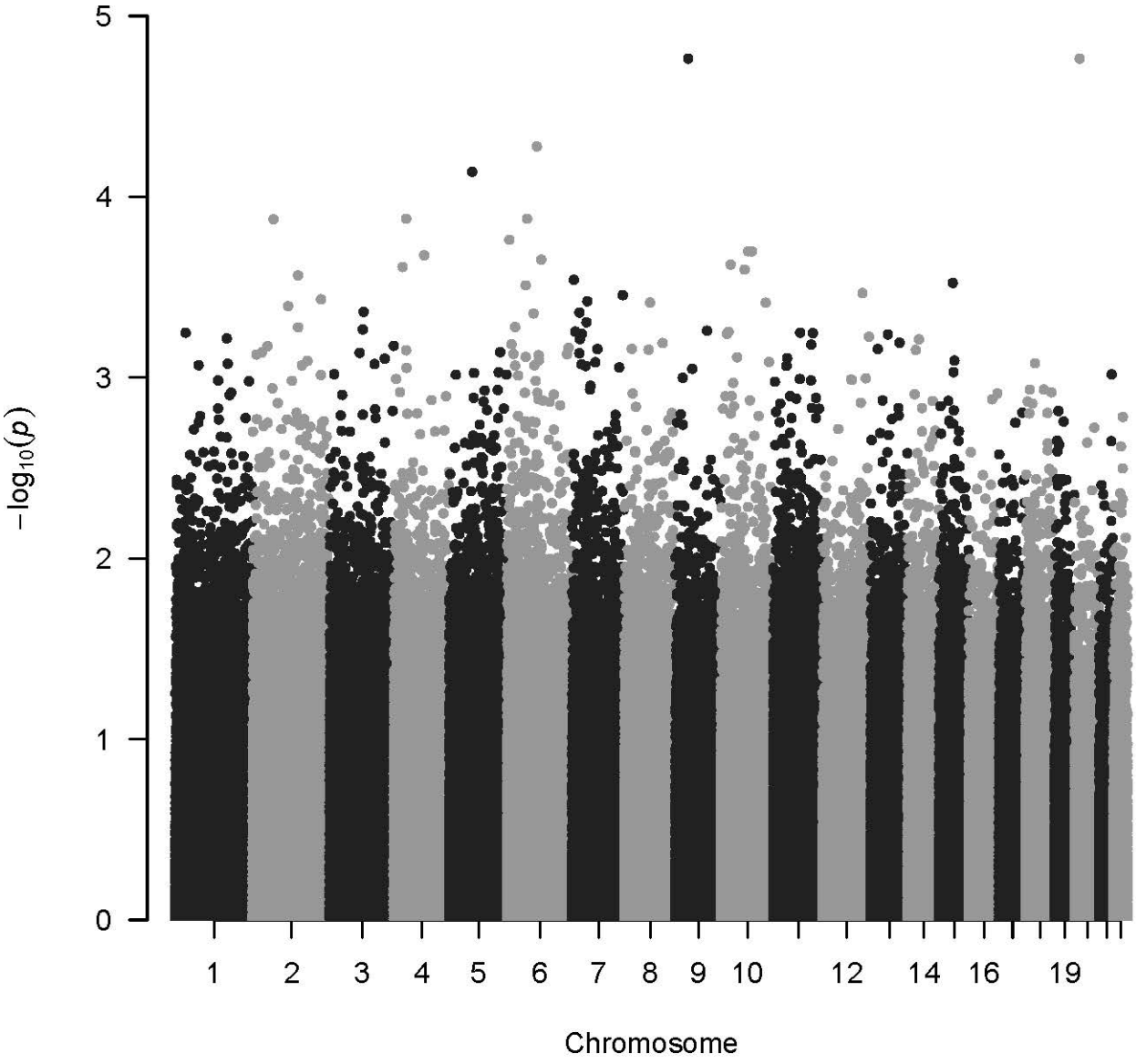
Supplementary Figure S9. Type I error rate and power of LIME$_{DSP}$ under 8 disease models and scenario 6 as given in Table 2. Three rows represent three data types: D, $D+1$ and $D+2$. The bars of color white, red and green refer to association, imprinting effect and maternal effect. The horizontal line marks the nominal a level of 0.05.



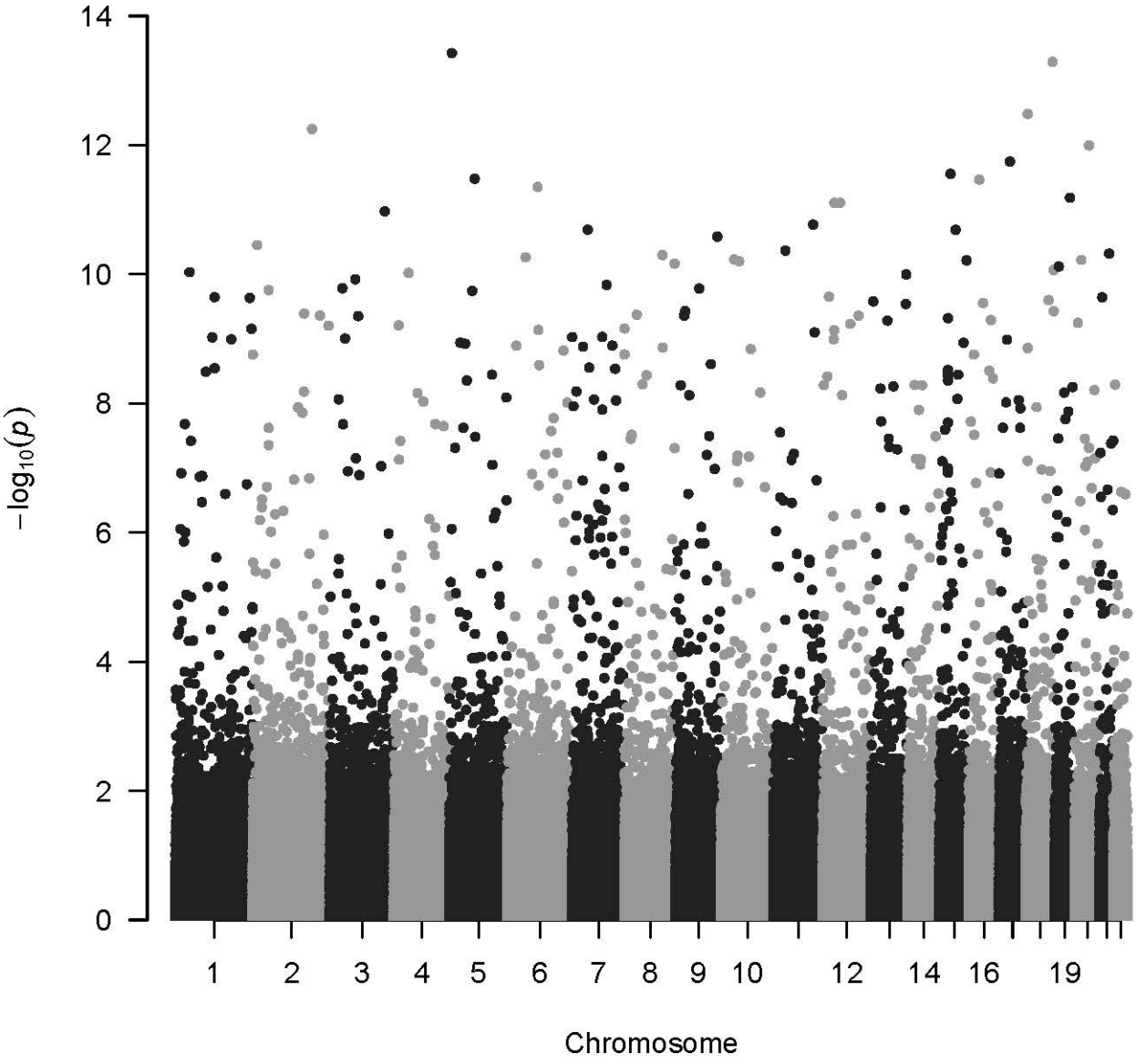
Supplementary Figure S10. Type I error rate and power of LIME$_{DSP}$ under 8 disease models and scenario 7 as given in Table 2. Three rows represent three data types: D, $D+1$ and $D+2$. The bars of color white, red and green refer to association, imprinting effect and maternal effect. The horizontal line marks the nominal a level of 0.05.



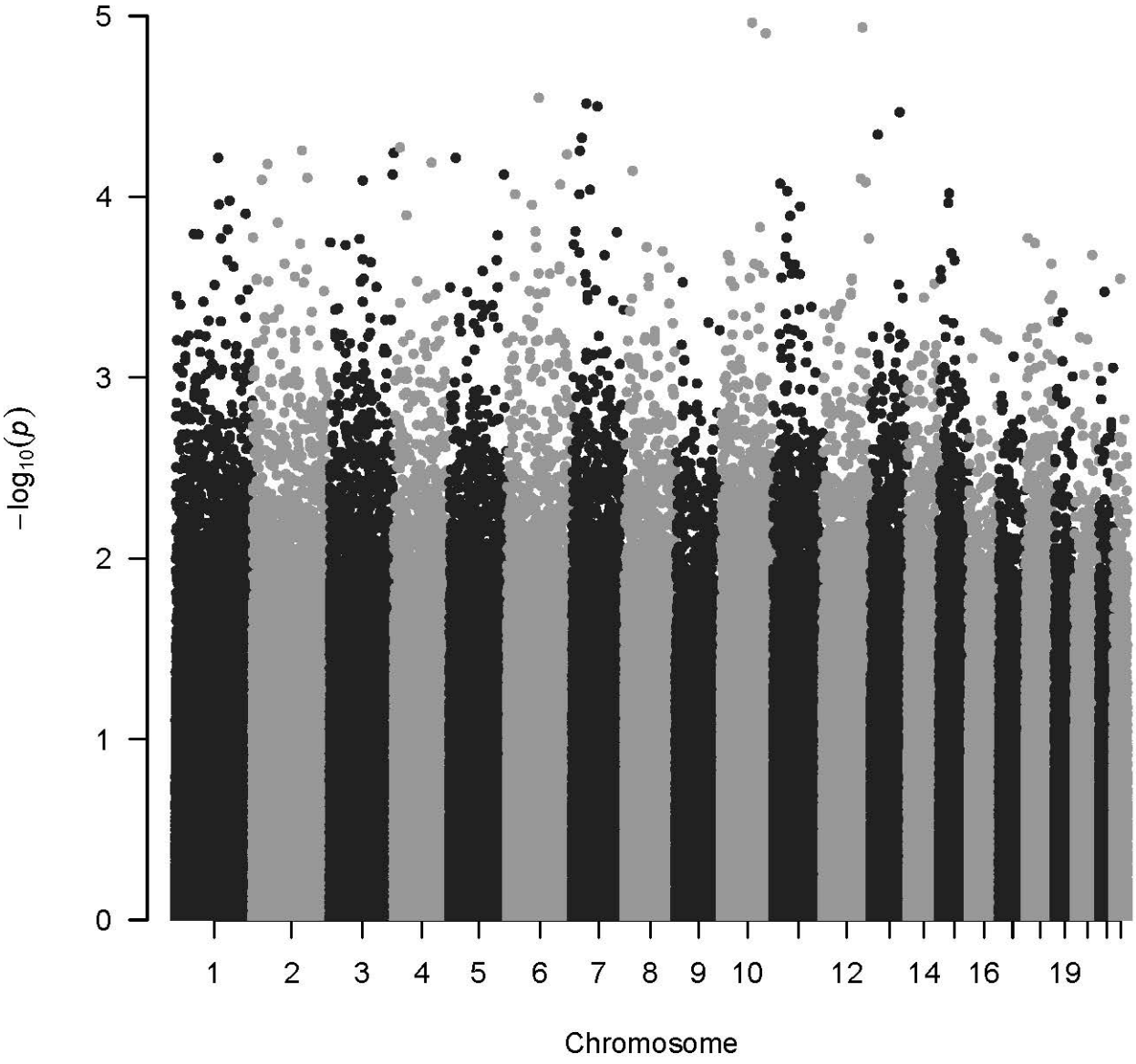
Supplementary Figure S11. Type I error rate and power of LIME$_{DSP}$ under 8 disease models and scenario 8 as given in Table 2. Three rows represent three data types: D, $D+1$ and $D+2$. The bars of color white, red and green refer to association, imprinting effect and maternal effect. The horizontal line marks the nominal a level of 0.05.



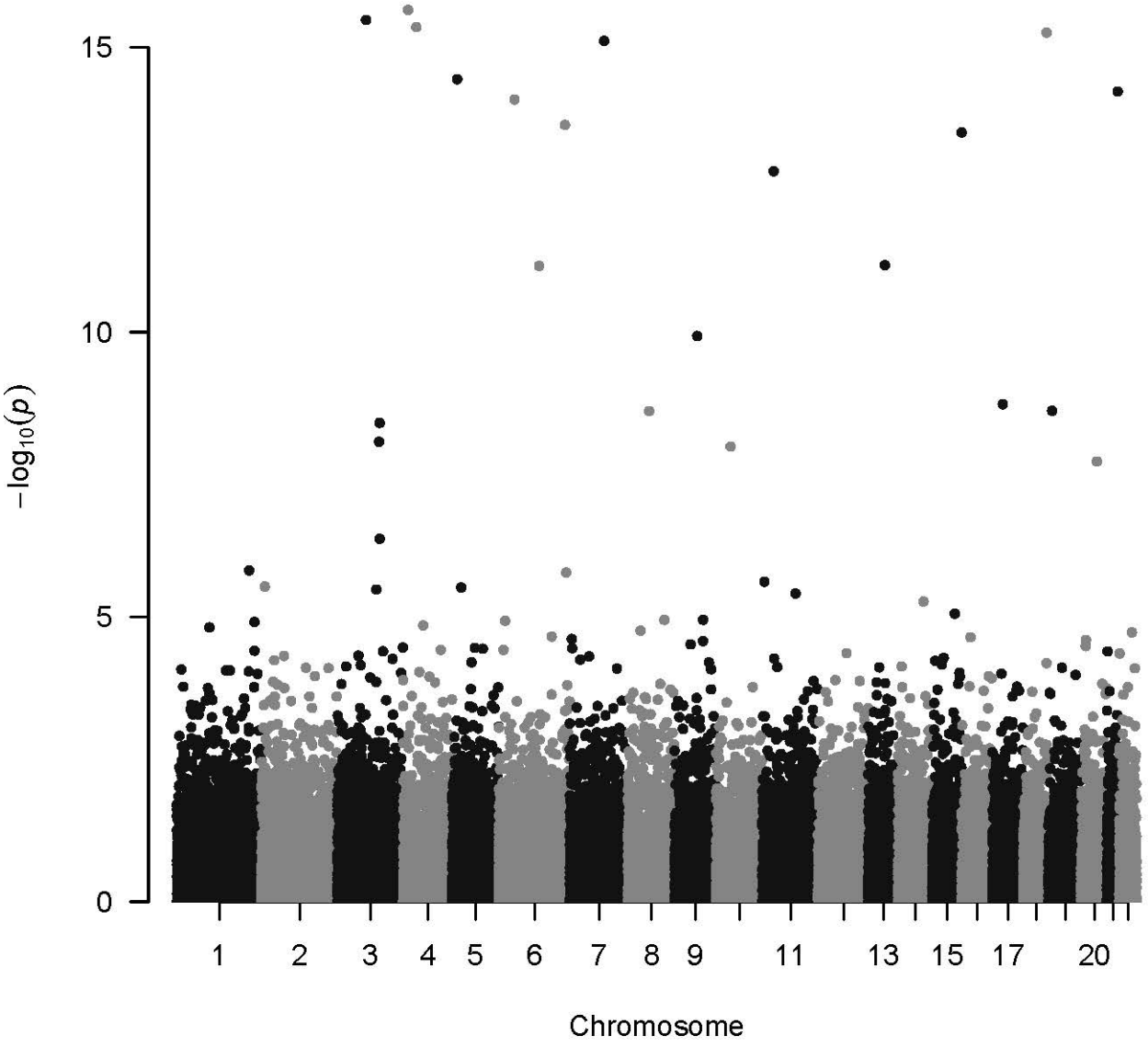
Supplementary Figure S12. Manhattan plot of -$\log_{10}$(p-value) for tests of association effect on club foot.



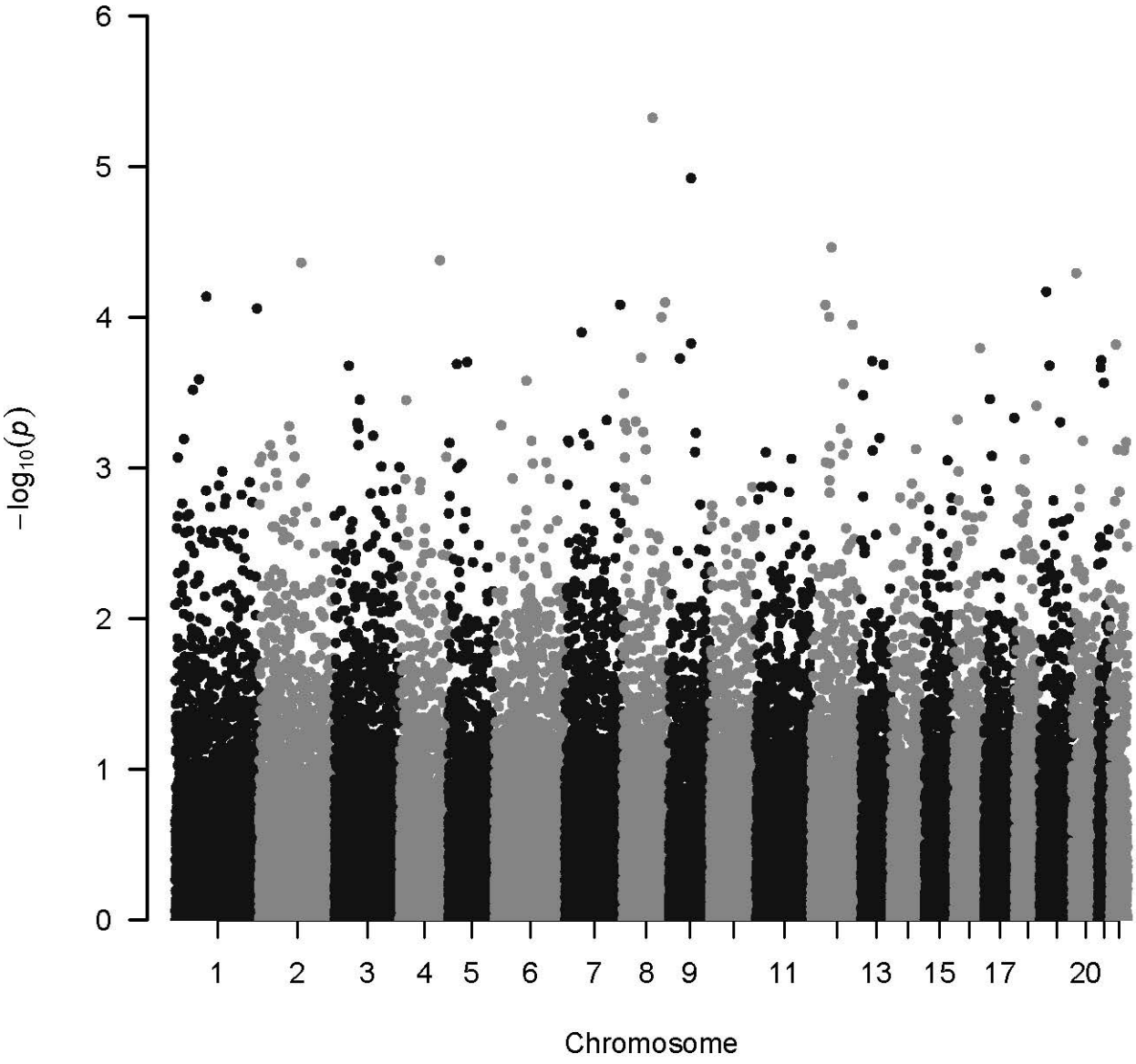
Supplementary Figure S13. Manhattan plot of -$\log_{10}$(p-value) for tests of imprinting effect on club foot.



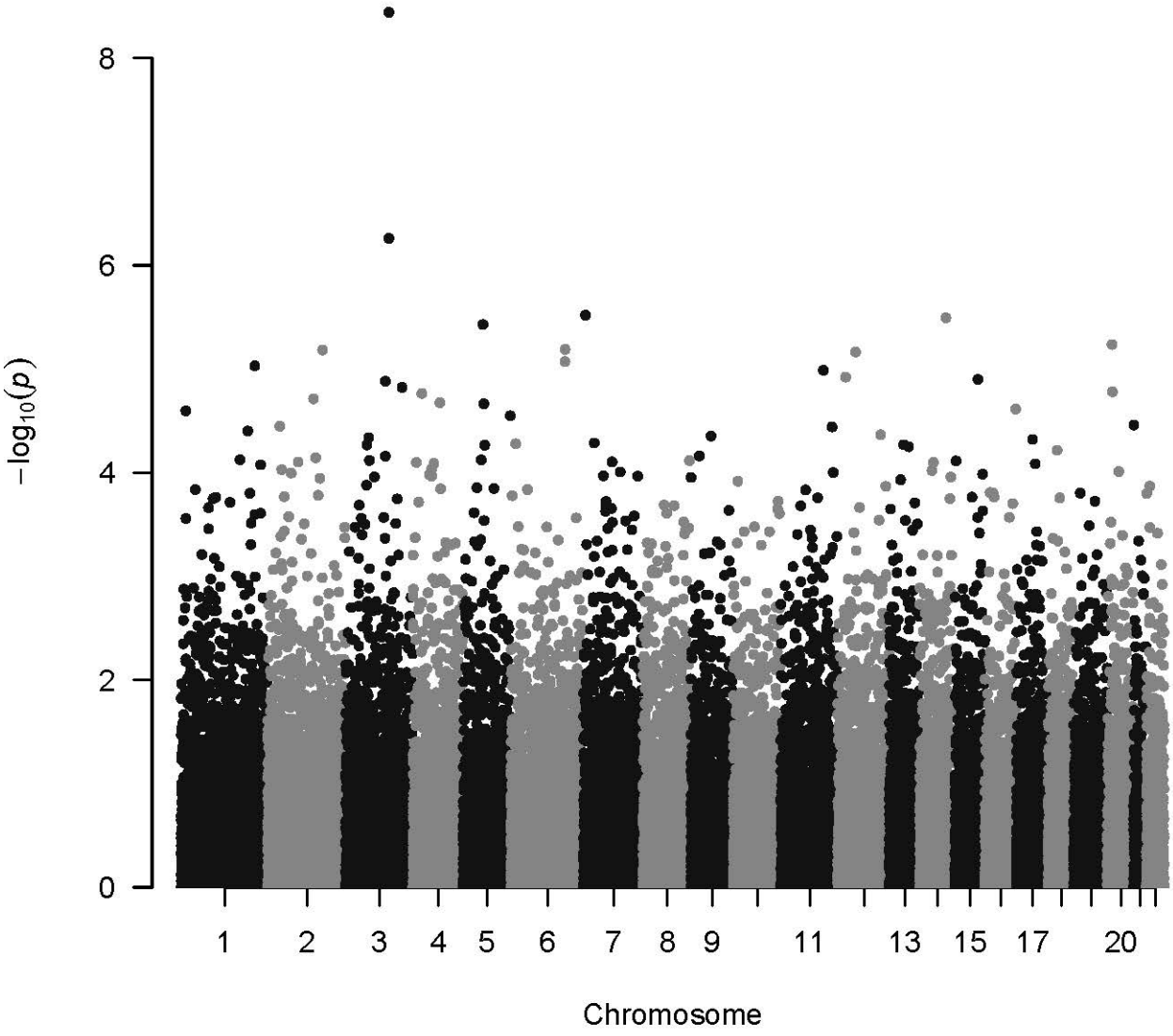
Supplementary Figure S14. Manhattan plot of -$\log_{10}$(p-value) for tests of maternal effect on club foot.



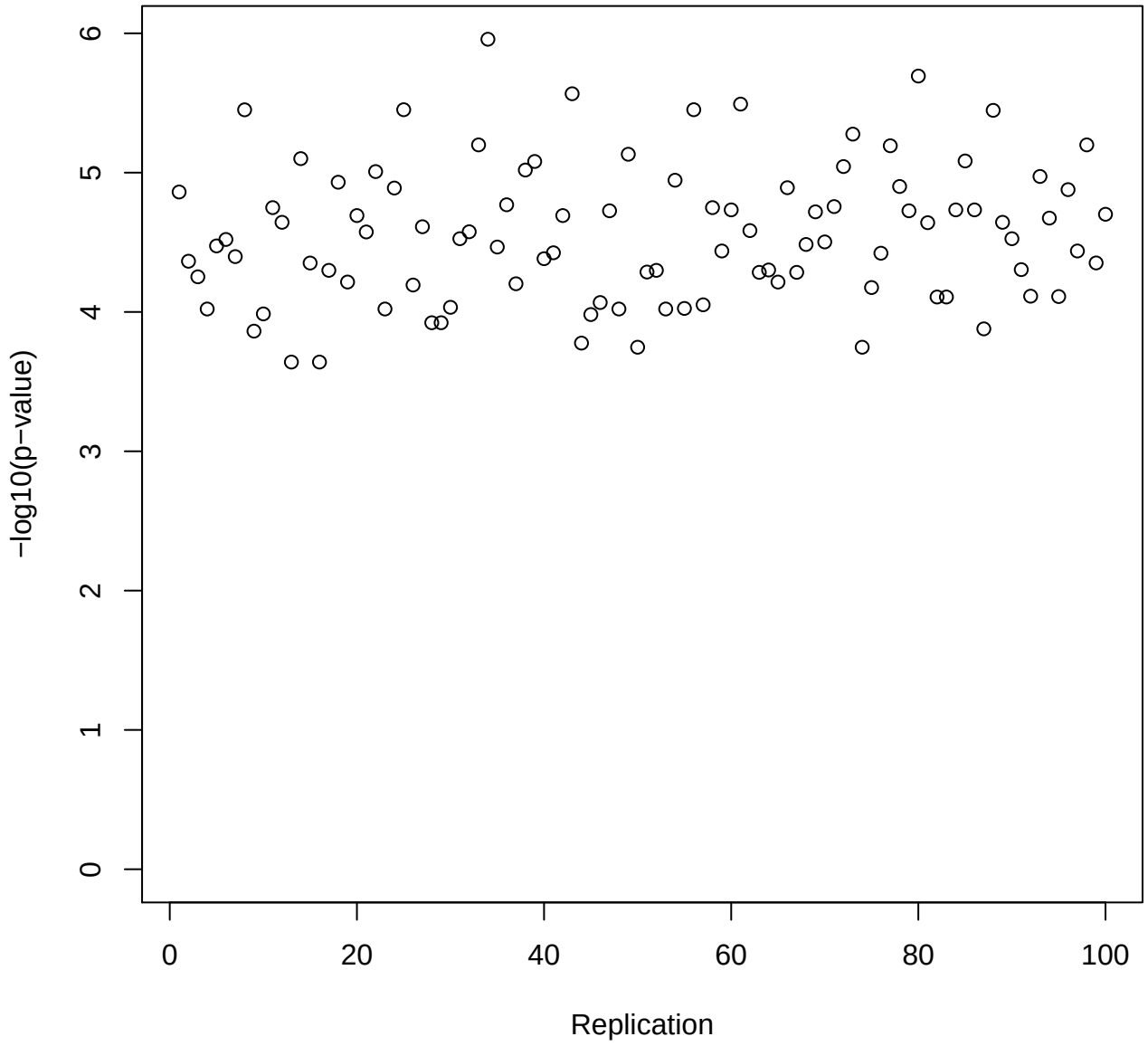
Supplementary Figure S15. Manhattan plot of -$\log_{10}$(p-value) for tests of association effect on FHS.



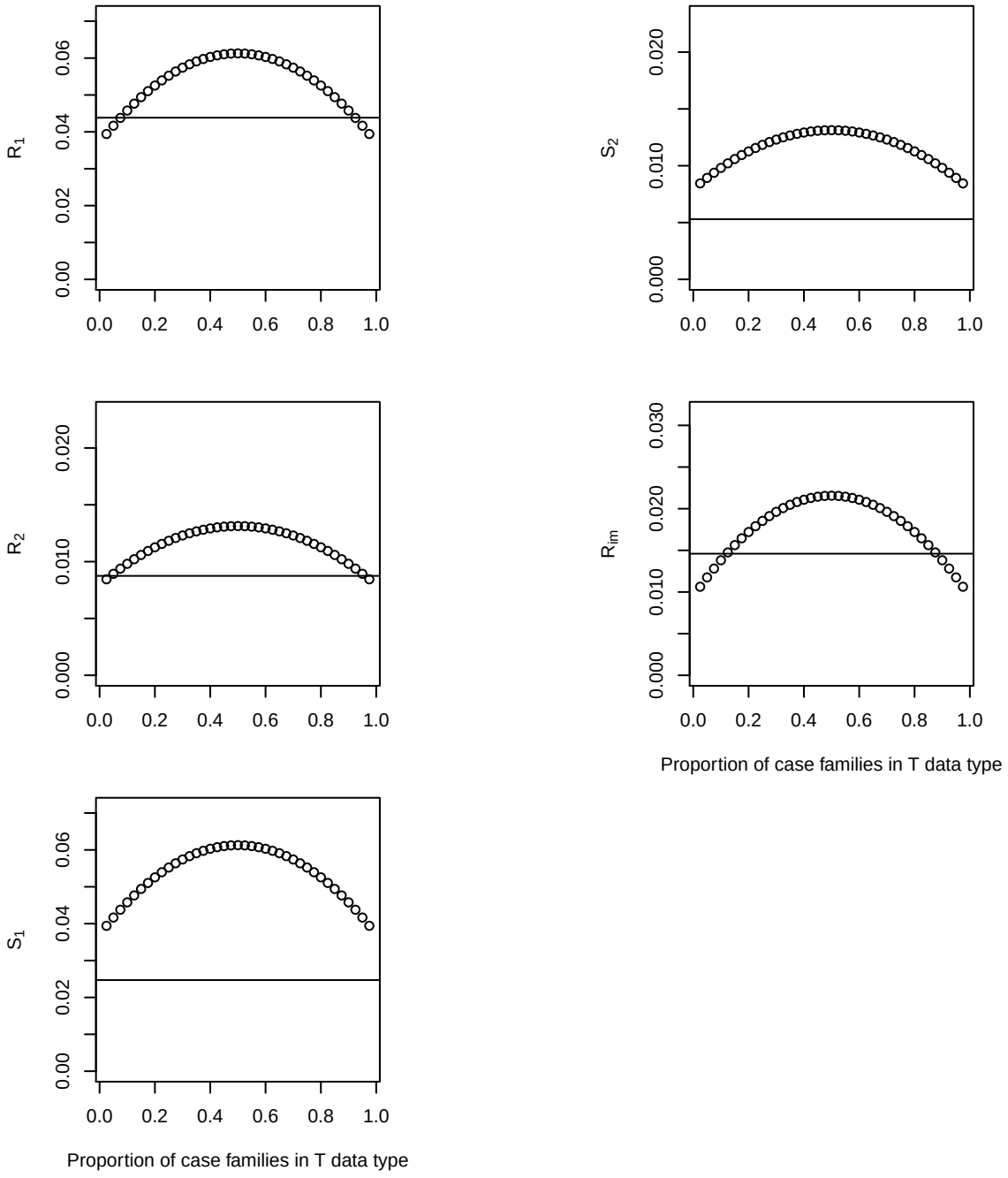
Supplementary Figure S16. Manhattan plot of -$\log_{10}$(p-value) for tests of imprinting effect on FHS.



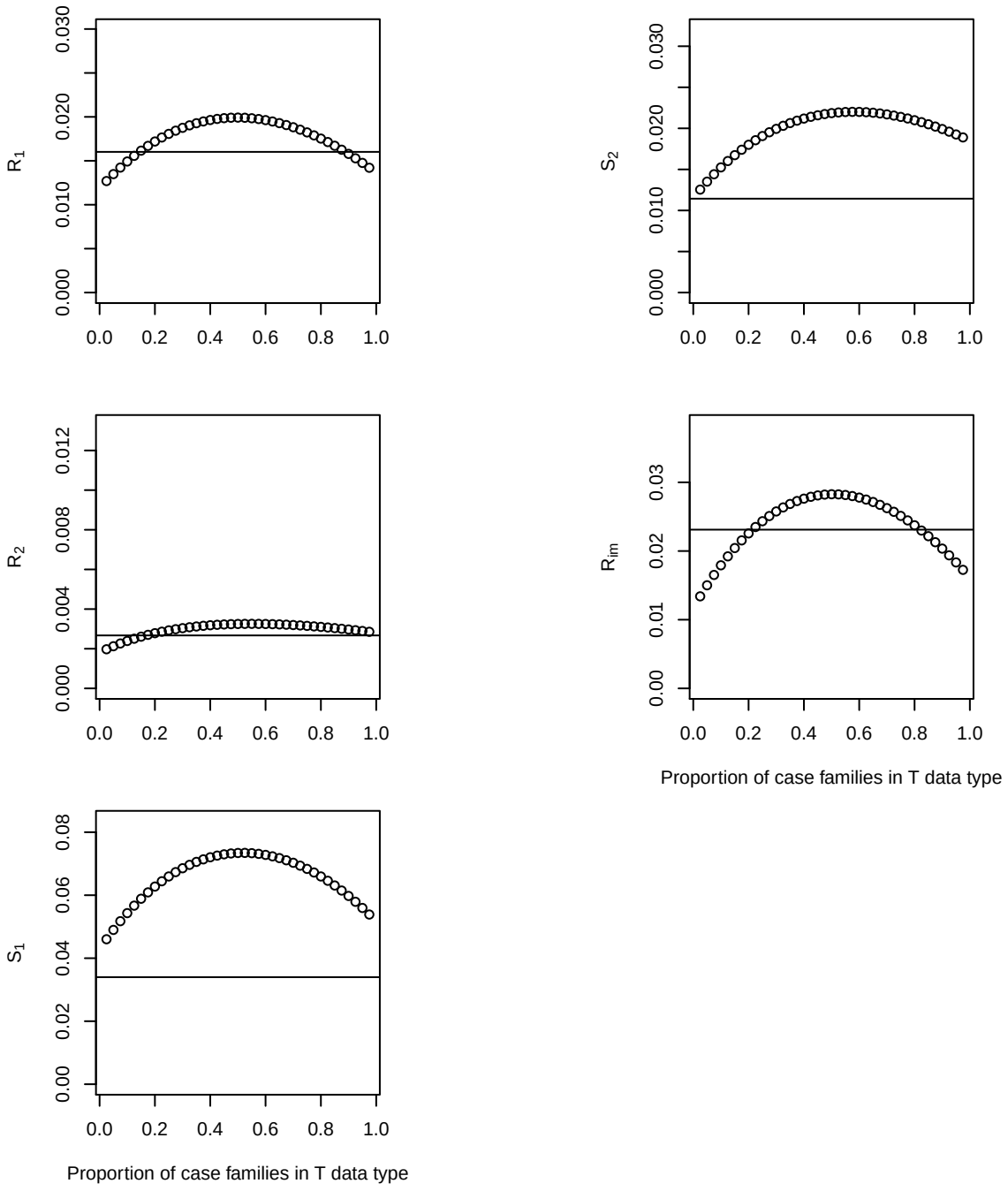
Supplementary Figure S17. Manhattan plot of -$\log_{10}$(p-value) for tests of maternal effect on FHS.



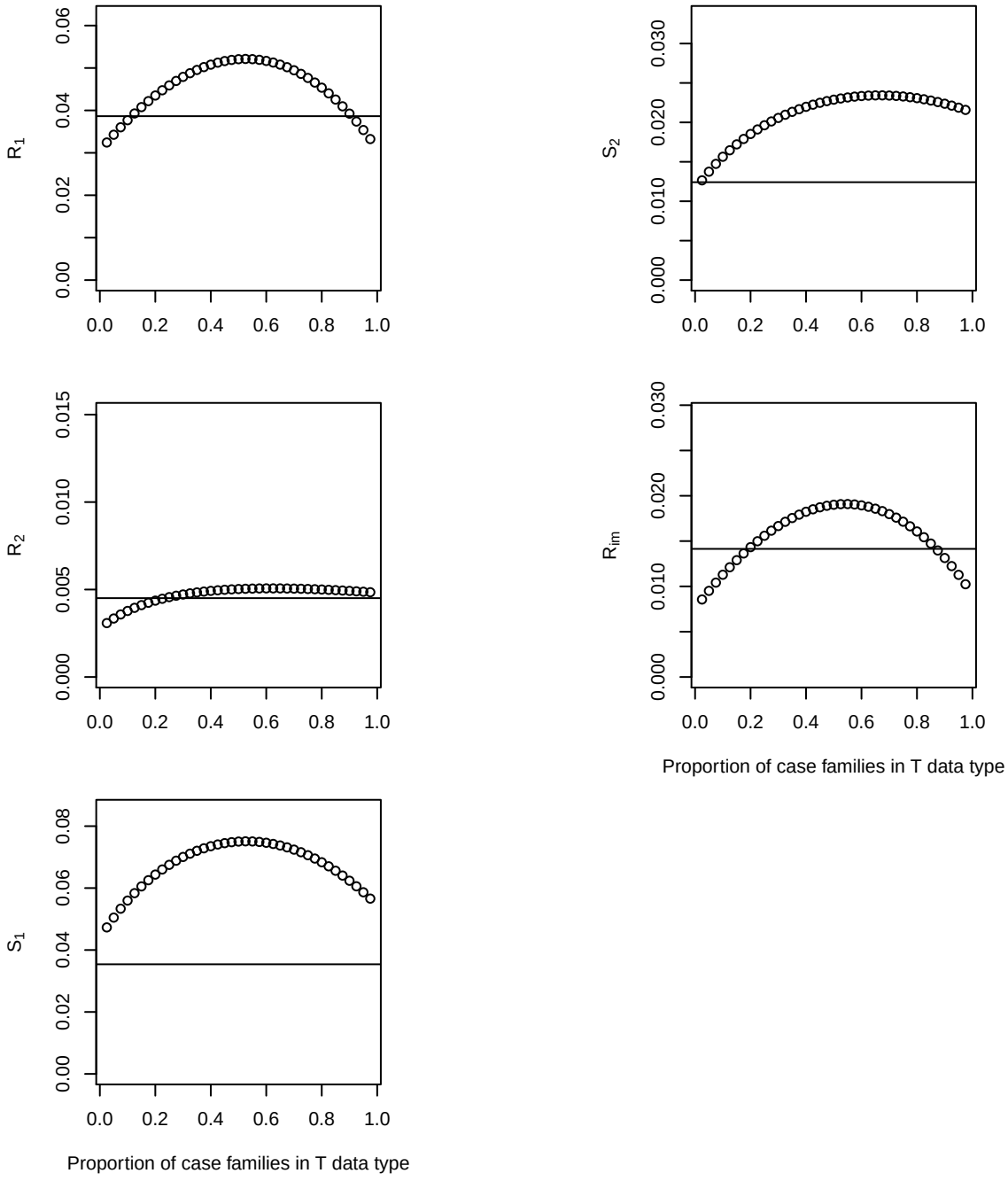
Supplementary Figure S18. Plot of the $-log_{10}$(p-values) for the imprinting effect of SNP rs1562705 versus replication index for proband designations from the FHS data.



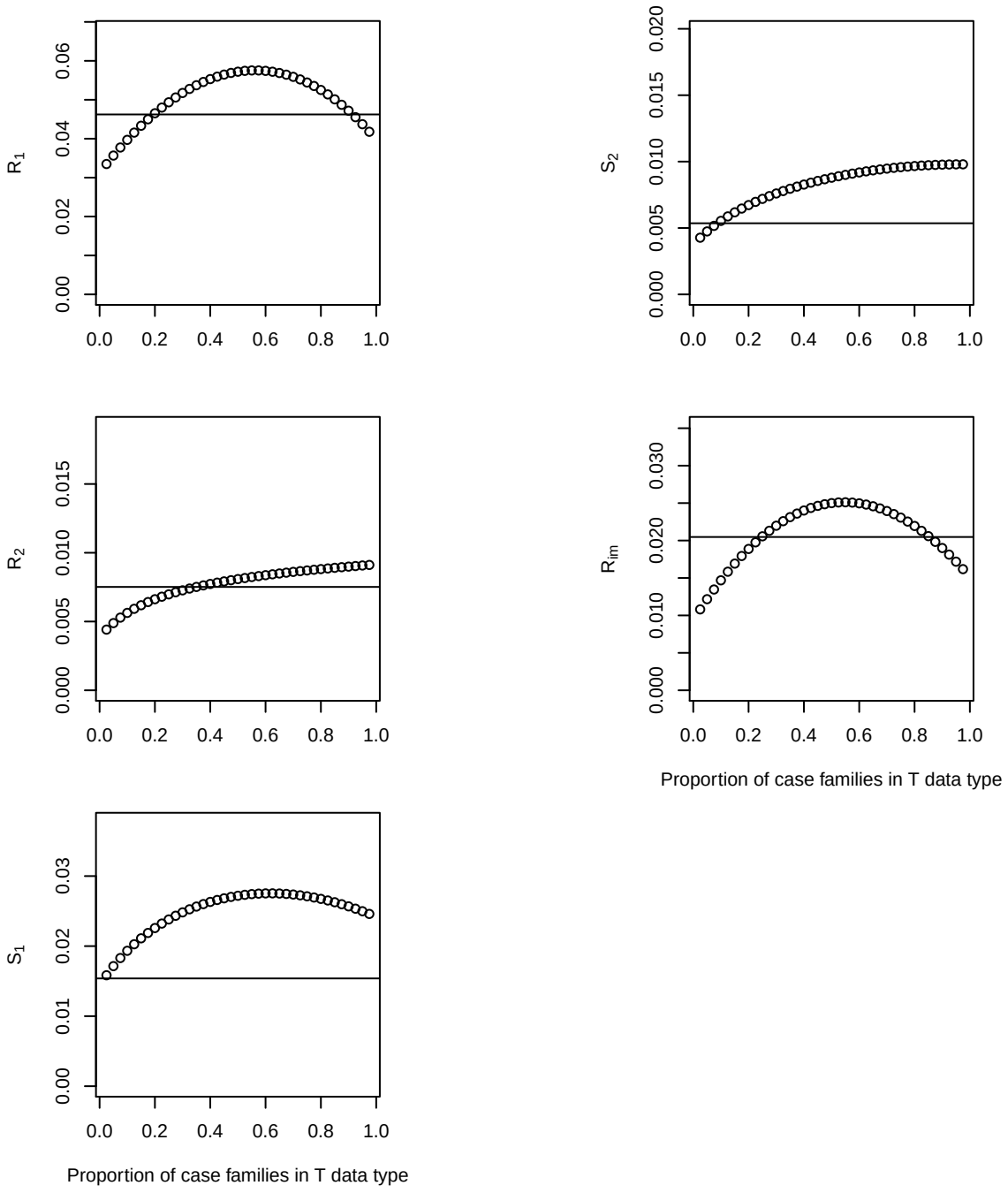
Supplementary Figure S19. Information content per individual for inference of parameters under disease model 1 and scenario 8. The horizontal line refers the information content per individual for LIME$_{DSP}$ applying to the D+2 design. The small circles represent information content per individual for LIME when applied to the T+3 design, with the proportion of case families varying from 0.025 to 0.975 by 0.025.



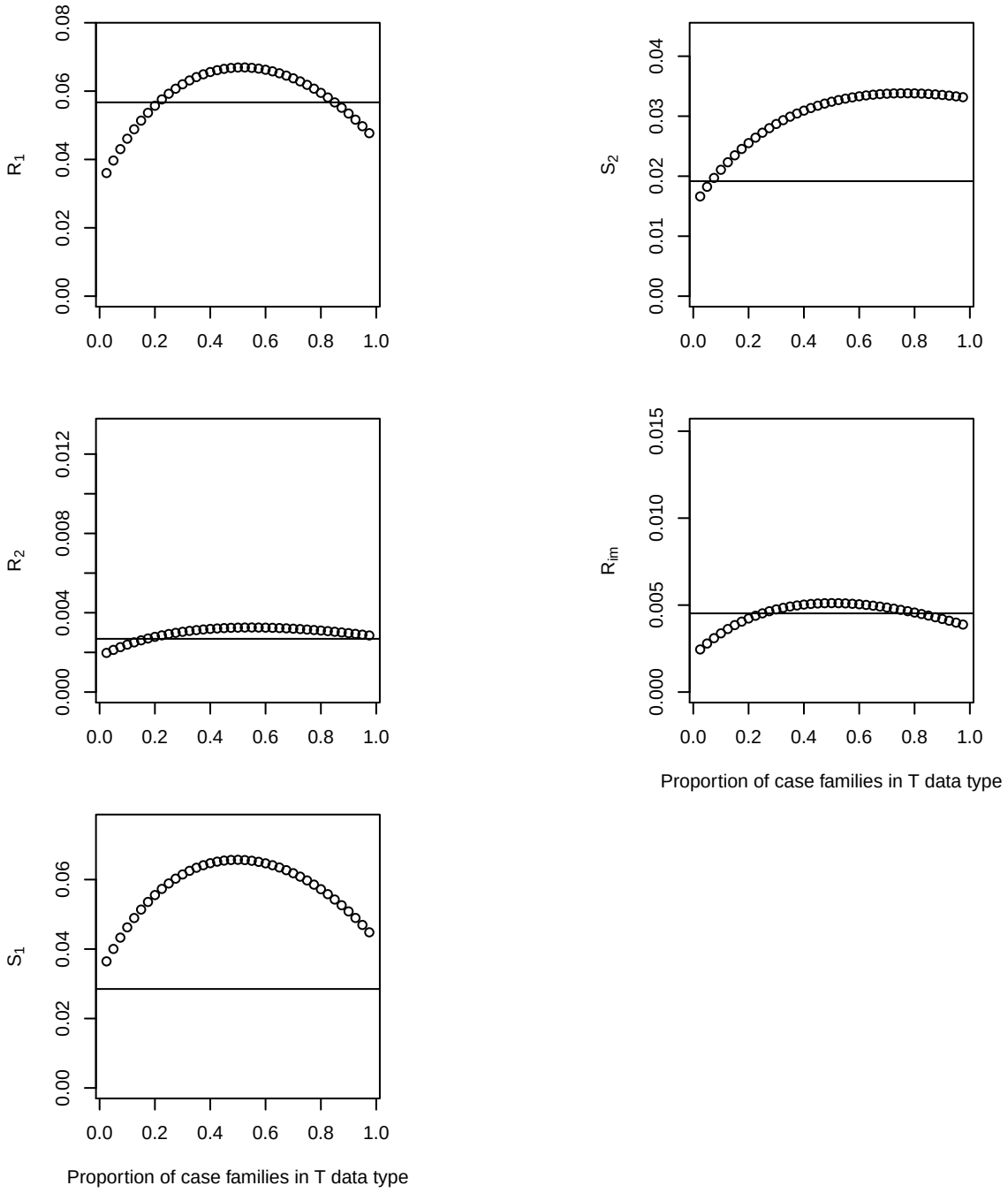
Supplementary Figure S20. Information content per individual for inference of parameters under disease model 2 and scenario 8. The horizontal line refers the information content per individual for LIME$_{DSP}$ applying to the D+2 design. The small circles represent information content per individual for LIME when applied to the T+3 design, with the proportion of case families varying from 0.025 to 0.975 by 0.025.



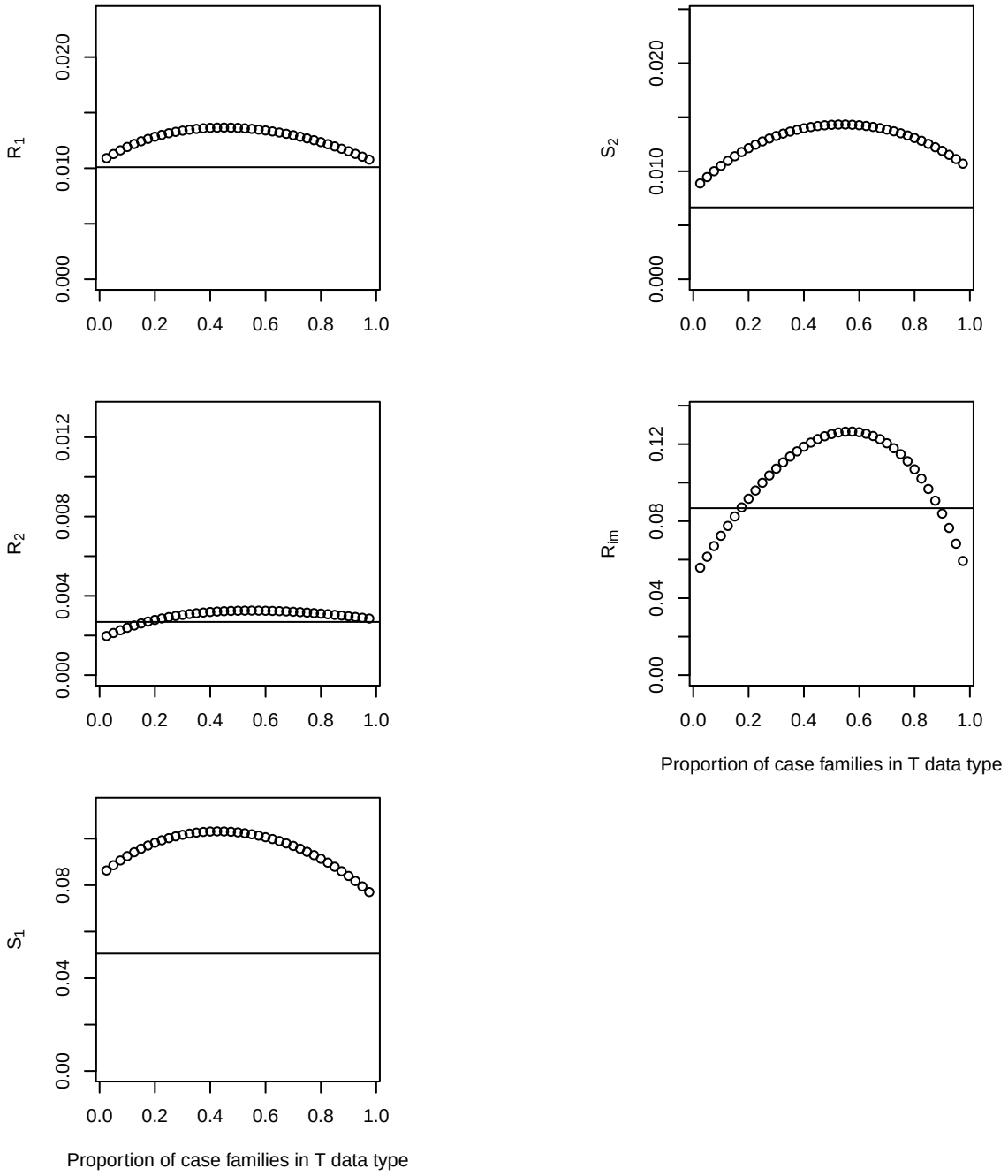
Supplementary Figure S21. Information content per individual for inference of parameters under disease model 3 and scenario 8. The horizontal line refers the information content per individual for LIME$_{DSP}$ applying to the D+2 design. The small circles represent information content per individual for LIME when applied to the T+3 design, with the proportion of case families varying from 0.025 to 0.975 by 0.025.



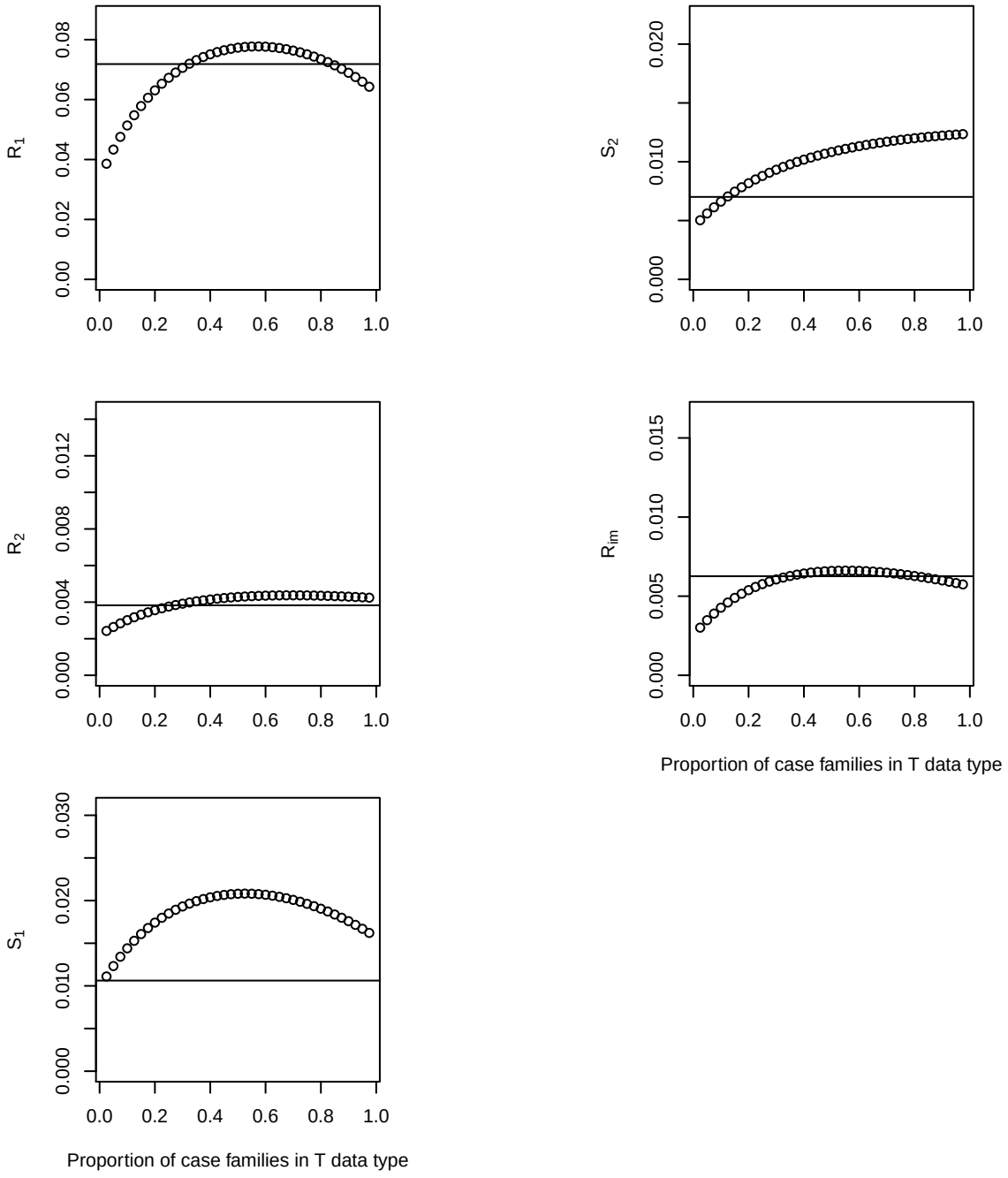
Supplementary Figure S22. Information content per individual for inference of parameters under disease model 4 and scenario 8. The horizontal line refers the information content per individual for LIME$_{DSP}$ applying to the D+2 design. The small circles represent information content per individual for LIME when applied to the T+3 design, with the proportion of case families varying from 0.025 to 0.975 by 0.025.



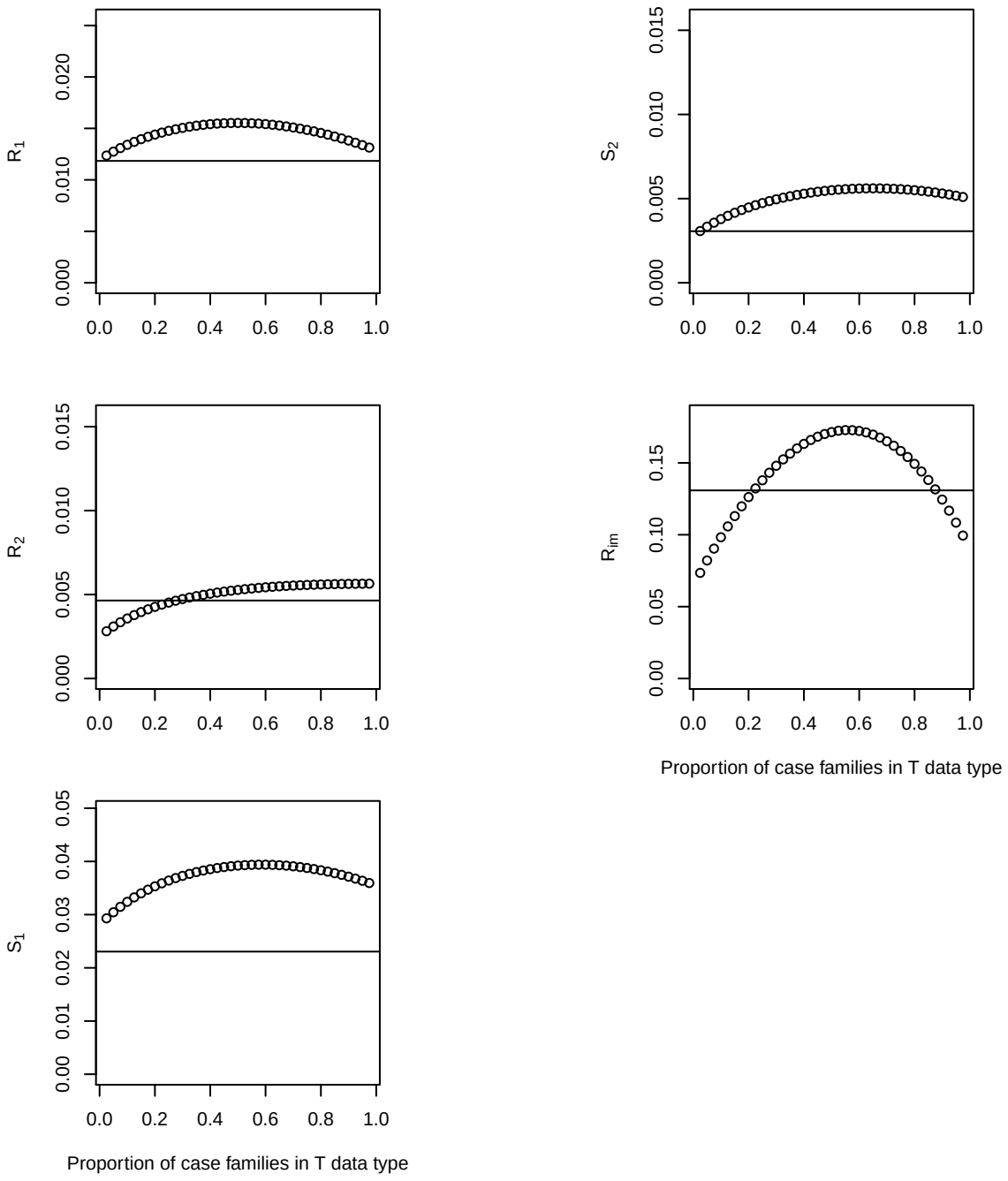
Supplementary Figure S23. Information content per individual for inference of parameters under disease model 5 and scenario 8. The horizontal line refers the information content per individual for LIME$_{DSP}$ applying to the D+2 design. The small circles represent information content per individual for LIME when applied to the T+3 design, with the proportion of case families varying from 0.025 to 0.975 by 0.025.



Supplementary Figure S24. Information content per individual for inference of parameters under disease model 6 and scenario 8. The horizontal line refers the information content per individual for LIME$_{DSP}$ applying to the D+2 design. The small circles represent information content per individual for LIME when applied to the T+3 design, with the proportion of case families varying from 0.025 to 0.975 by 0.025.



Supplementary Figure S25. Information content per individual for inference of parameters under disease model 7 and scenario 8. The horizontal line refers the information content per individual for LIME$_{DSP}$ applying to the D+2 design. The small circles represent information content per individual for LIME when applied to the T+3 design, with the proportion of case families varying from 0.025 to 0.975 by 0.025.



Supplementary Figure S26. Information content per individual for inference of parameters under disease model 8 and scenario 8. The horizontal line refers the information content per individual for LIME$_{DSP}$ applying to the D+2 design. The small circles represent information content per individual for LIME when applied to the T+3 design, with the proportion of case families varying from 0.025 to 0.975 by 0.025.