Supplementary material for "Adjusting for non-confounding covariates in case-control association studies"

Siliang Zhang[1], Jinbo Chen[2], Zhiliang Ying[3], Hong Zhang[4]

[1] *Key Laboratory of Advanced Theory and Application in Statistics and Data Science, Ministry of Education, School of Statistics, East China Normal University, Shanghai 200062, China*

[2] *Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA*

[3] *Department of Statistics, Columbia University, New York, NY 10027, USA*

[4] *Department of Statistics and Finance, School of Management, University of Science and Technology of China, Hefei, Anhui 230026, China*

The proofs of Lemma 1 and Corollary 1 are presented in Section S1 and Section S2, respectively. The proofs of Lemma 2 and Corollary 2 are given in Section S3. The definition of Pitman's asymptotic relative efficiency (Section 3.3 of the main text) is restated in Section S4. The proofs of Theorem 1 and Corollary 3 are presented in Section S5. The proof of Theorem 2 is presented in Section S6. The proof of Theorem 3 is given in Section S7.

S1 Proof of Lemma 1

We adopt the notations of the main text, for example,

$$f = \mathrm{pr}(D = 1), \quad \theta = \mathrm{pr}(X = 1), \quad \pi = \mathrm{pr}(E = 1).$$

We also introduce some additional notations:

$$p_i = \mathrm{pr}(D = 1 \mid E = i), \quad q_i = 1 - p_i = \mathrm{pr}(D = 0 \mid E = i),$$

$$p_{ij} = \mathrm{pr}(D = 1 \mid X = i, E = j), \quad q_{ij} = 1 - p_{ij} = \mathrm{pr}(D = 0 \mid X = i, E = j), \tag{S1.1}$$

for $i = 0, 1;\ j = 0, 1$.

Throughout this document, we assume that X and E are independent unless specially noted, so that $p_i = p_{1i}\theta + p_{0i}(1 - \theta),\ q_i = q_{1i}\theta + q_{0i}(1 - \theta)$. Under the retrospective setting, the random variables n_{1+1} and n_{0+1} follow binomial distributions, i.e., $n_{1+1} \sim B(n_{1++}, p'_1)$ and $n_{0+1} \sim B(n_{0++}, p'_0)$, where

$$p'_1 = \mathrm{pr}(E = 1 \mid D = 1) = \frac{p_1\pi}{f}, \quad p'_0 = \mathrm{pr}(E = 1 \mid D = 0) = \frac{q_1\pi}{1 - f}, \tag{S1.2}$$

$$q'_1 = \mathrm{pr}(E = 0 \mid D = 1) = \frac{p_0(1 - \pi)}{f}, \quad q'_0 = \mathrm{pr}(E = 0 \mid D = 0) = \frac{q_0(1 - \pi)}{1 - f}. \tag{S1.3}$$

For any $f \in (0, 1)$, it follows from the standard large sample theory for the sample odds ratio and (S1.1)-(S1.3) that

$$\begin{aligned}
\hat{\gamma}_M &= \log \frac{p_1' q_0'}{p_0' q_1'} + O_P(n^{-\frac{1}{2}}) \\
&= \log \frac{p_1 q_0}{p_0 q_1} + O_P(n^{-\frac{1}{2}}) \\
&= \log \frac{\{p_{11}\theta + p_{01}(1-\theta)\}\{q_{10}\theta + q_{00}(1-\theta)\}}{(p_{10}\theta + p_{00}(1-\theta))\,(q_{11}\theta + q_{01}(1-\theta))} + O_P(n^{-\frac{1}{2}}) \\
&= \gamma + \log \left\{ \frac{(1-\theta+e^{\alpha+\beta+\gamma}+e^{\beta}\theta)(1+e^{\alpha}\theta+e^{\alpha+\beta}(1-\theta))}{(1-\theta+e^{\alpha+\beta}+e^{\beta}\theta)(1+e^{\alpha+\gamma}\theta+e^{\alpha+\beta+\gamma}(1-\theta))} \right\} + O_P(n^{-\frac{1}{2}}) \\
&= \gamma + \log \left\{ 1 + \frac{e^{\alpha}(b_1-b_2)(1-e^{\gamma})}{(1+e^{\alpha}b_2)(1+e^{\alpha+\gamma}b_1)} \right\} + O_P(n^{-\frac{1}{2}}) \\
&= \gamma + \delta + O_P(n^{-\frac{1}{2}}), \qquad \text{(S1.4)}
\end{aligned}$$

where

$$b_1 = e^{\beta}(1-\theta)+\theta = 1+(e^{\beta}-1)(1-\theta),\; b_2 = e^{\beta}/(1-\theta+e^{\beta}\theta) = \frac{1}{1+(e^{-\beta}-1)(1-\theta)},$$

$$\hat{\gamma}_M = \log(n_{1+1}/n_{1+0}) - \log(n_{0+1}/n_{0+0}),$$

and

$$\delta = \log \left\{ 1 + \frac{e^{\alpha}(b_1-b_2)(1-e^{\gamma})}{(1+e^{\alpha+\gamma}b_1)(1+e^{\alpha}b_2)} \right\}. \qquad \text{(S1.5)}$$

Moreover,

$$b_1 = 1+(e^{\beta}-1)(1-\theta) = \frac{1+(e^{\beta}+e^{-\beta}-2)\theta(1-\theta)}{1+(e^{-\beta}-1)(1-\theta)} \geq \frac{1}{1+(e^{-\beta}-1)(1-\theta)} = b_2 > 0. \qquad \text{(S1.6)}$$

S2 Proof of Corollary 1

It follows from $b_1 \geq b_2 > 0$ that

$$-\gamma < \delta \leq 0 \text{ if } \gamma > 0 \text{ and } 0 \leq \delta < -\gamma \text{ if } \gamma < 0 \text{ and } \delta = 0 \text{ if } \gamma = 0, \quad \text{(S2.1)}$$

so that

$$|\gamma + \delta| \leq |\gamma|.$$

Furthermore, it is easily seen from the expression of δ given in (S1.5) of the main text that $\delta = 0$ if and only if $b_1 = b_2$ (which leads to $\beta = 0$, i.e., X is not associated with D) or $\gamma = 0$ (i.e., E is not associated with D). Finally, setting the derivative of (S1.5) with respect to α to be 0, we can see that $|\delta|$ is minimized at $\alpha_{\min}$ defined in (3.8) of the main text.

S3 Proofs of Lemma 2 and Corollary 2

As defined in the main text, $\nu = n_{1++}/n_{0++}$, so that $n_{0++} = n/(1+\nu)$ and $n_{1++} = n\nu/(1+\nu)$. Assume a contiguous alternative scenario where $\gamma = cn^{-1/2}$.

First, we derive the asymptotic distribution of $\hat{\gamma}_M$. According to the standard large sample theory, the regularity conditions R1 - R3 (see Chapter 4, Serfling, 2009) hold for logistic regression models, which gives

$$n^{1/2}(\hat{\gamma}_M - \gamma - \delta) \rightarrow N(0, \sigma_M^2) \text{ in distribution,}$$

where the asymptotic variance is

$$\sigma_M^2 = \frac{n}{n_{0++}p_0'q_0'} + \frac{n}{n_{1++}p_1'q_1'} = \frac{(1+\nu)}{p_0'q_0'} + \frac{(1+\nu)}{\nu p_1'q_1'}. \tag{S3.1}$$

Since $\gamma = 0$ implies that $p_1' = p_0' = \pi$, we have that

$$\sigma_M^2 \to \sigma_0^2 \text{ as } \gamma \to 0, \tag{S3.2}$$

where $\sigma_0^2 = (2 + \nu + 1/\nu)/\{\pi(1-\pi)\}$. The above results hold for any $f \in (0,1)$.

Next we derive the asymptotic distribution of $\hat{\gamma}_A$. According to Gart (1962), we have

$$n^{1/2}\left(\hat{\gamma}_A - \gamma\right) \to N(0, \sigma_A^2) \text{ in distribution,}$$

where

$$\begin{aligned}\sigma_A^2 =& \left\{ \left(\frac{n}{n_{0++}d_{00}h_{00}(1-h_{00})} + \frac{n}{n_{1++}d_{01}h_{01}(1-h_{01})} \right)^{-1} + \right. \\ & \left. \left(\frac{n}{n_{0++}d_{10}h_{10}(1-h_{10})} + \frac{n}{n_{1++}d_{11}h_{11}(1-h_{11})} \right)^{-1} \right\}^{-1} \\ =& \left\{ \left(\frac{1+\nu}{d_{00}h_{00}(1-h_{00})} + \frac{1+\nu}{\nu d_{01}h_{01}(1-h_{01})} \right)^{-1} + \right. \\ & \left. \left(\frac{1+\nu}{d_{10}h_{10}(1-h_{10})} + \frac{1+\nu}{\nu d_{11}h_{11}(1-h_{11})} \right)^{-1} \right\}^{-1}\end{aligned} \tag{S3.3}$$

and

$$d_{ij} = \text{pr}(X = i \mid D = j), \quad h_{ij} = \text{pr}(E = 1 \mid X = i, D = j).$$

If we denote

$$a_{10} = \frac{d_{00}h_{00}}{1+\nu}, \quad a_{20} = \frac{d_{00}(1-h_{00})}{1+\nu}, \quad a_{30} = \frac{\nu d_{01}h_{01}}{1+\nu}, \quad a_{40} = \frac{\nu d_{01}(1-h_{01})}{1+\nu},$$
$$a_{11} = \frac{d_{10}h_{10}}{1+\nu}, \quad a_{21} = \frac{d_{10}(1-h_{10})}{1+\nu}, \quad a_{31} = \frac{\nu d_{11}h_{11}}{1+\nu}, \quad a_{41} = \frac{\nu d_{11}(1-h_{11})}{1+\nu},$$

then we have

$$\sigma_M^2 = \frac{1}{a_{10}+a_{11}} + \frac{1}{a_{20}+a_{21}} + \frac{1}{a_{30}+a_{31}} + \frac{1}{a_{40}+a_{41}}$$

and

$$\sigma_A^2 = \left\{ \left(\frac{1}{a_{10}} + \frac{1}{a_{20}} + \frac{1}{a_{30}} + \frac{1}{a_{40}}\right)^{-1} + \left(\frac{1}{a_{11}} + \frac{1}{a_{21}} + \frac{1}{a_{31}} + \frac{1}{a_{41}}\right)^{-1} \right\}^{-1}.$$

Applying the Minkowski inequality, we immediately have that $\sigma_M^2 \leq \sigma_A^2$, and the inequality holds even when X and E are correlated. Moreover, the equality holds if and only if $a_{i1} = ka_{i0}$ $(i = 1, \ldots, 4)$, or equivalently, X is independent of D (i.e., $\beta = 0$).

Next, we compare σ_M^2 and σ_A^2 under the condition of $\gamma \to 0$. We rewrite the asymptotic variances as

$$\begin{aligned}
\sigma_M^2 =& \frac{(1+\nu)(1-f)^2}{q_0q_1\pi(1-\pi)} + \frac{(1+\nu)f^2}{\nu p_0p_1\pi(1-\pi)} \\
=& \frac{1}{\pi(1-\pi)}\left\{\frac{(1+\nu)(1-f)^2}{q_0q_1} + \frac{(1+\nu)f^2}{\nu p_0p_1}\right\} \\
=& \frac{1+\nu}{\pi(1-\pi)}\left\{\frac{(1-f)^2}{E(q_{X0})E(q_{X1})} + \frac{f^2}{\nu E(p_{X0})E(p_{X1})}\right\}
\end{aligned}$$

and

$$
\begin{aligned}
\sigma_A^2 =& \Bigg\{(1-\theta)\left(\frac{(1+\nu)(1-f)}{q_{01}\pi}+\frac{(1+\nu)(1-f)}{q_{00}(1-\pi)}+\frac{(1+\nu)f}{\nu p_{01}\pi}+\frac{(1+\nu)f}{\nu p_{00}(1-\pi)}\right)^{-1}+ \\
&\theta\left(\frac{(1+\nu)(1-f)}{q_{11}\pi}+\frac{(1+\nu)(1-f)}{q_{10}(1-\pi)}+\frac{(1+\nu)f}{\nu p_{11}\pi}+\frac{(1+\nu)f}{\nu p_{10}(1-\pi)}\right)^{-1}\Bigg\}^{-1} \\
=& (1+\nu)\left[E\left\{\frac{(1-f)}{q_{X1}\pi}+\frac{(1-f)}{q_{X0}(1-\pi)}+\frac{f}{\nu p_{X1}\pi}+\frac{f}{\nu p_{X0}(1-\pi)}\right\}^{-1}\right]^{-1}.
\end{aligned}
$$

If $\gamma \to 0$, then $p_{i1} \to p_{i0}$ and $q_{i1} \to q_{i0}$ for $i = 1, 2$. Consequently,

$$
\begin{aligned}
\lim_{\gamma\to 0}\frac{\sigma_A^2}{\sigma_M^2} =& \frac{\left[E\left\{\frac{(1-f)}{q_{X0}}+\frac{f}{\nu p_{X0}}\right\}^{-1}\right]^{-1}}{\left(\frac{1-f}{Eq_{X0}}\right)^2+\frac{1}{\nu}\left(\frac{f}{Ep_{X0}}\right)^2} \\
=& \frac{(1+\frac{1}{\nu})\left[\left\{1+(\frac{1-\theta}{\theta})(\frac{1+e^{\alpha+\beta}}{1+e^{\alpha}})(\frac{\nu+e^{-\beta}}{\nu+1})\right\}^{-1}+\left\{1+(\frac{\theta}{1-\theta})(\frac{1+e^{\alpha}}{1+e^{\alpha+\beta}})(\frac{\nu+e^{\beta}}{\nu+1})\right\}^{-1}\right]^{-1}}{1+\frac{1}{\nu}} \\
=& \left[\left\{1+\left(\frac{1-\theta}{\theta}\right)\left(\frac{1+e^{\alpha+\beta}}{1+e^{\alpha}}\right)\left(\frac{\nu+e^{-\beta}}{\nu+1}\right)\right\}^{-1}\right. \\
&\left.+\left\{1+\left(\frac{\theta}{1-\theta}\right)\left(\frac{1+e^{\alpha}}{1+e^{\alpha+\beta}}\right)\left(\frac{\nu+e^{\beta}}{\nu+1}\right)\right\}^{-1}\right]^{-1} \qquad \text{(S3.4)} \\
=& 1+\frac{\nu\theta(1-\theta)}{(1+\nu)}\frac{(1-e^{\beta})^2}{\left\{(1-\theta)\phi+e^{\beta}\theta\phi^{-1}\right\}^2+\nu e^{\beta}\left\{(1-\theta)\phi+\theta\phi^{-1}\right\}^2}, \\
=& \lambda, \qquad \text{(S3.5)}
\end{aligned}
$$

where $\phi = \sqrt{\frac{1+e^{\alpha+\beta}}{1+e^{\alpha}}}$ and $\lambda \geq 1$, and $\lambda = 1$ if and only if $\beta = 0$.

Denote $\rho = e^{\alpha}$. In the rare outcome case ($f \to 0$ or equivalently $\rho \to 0$), applying Taylor's expansion to (S3.4), we have

$$\begin{aligned}
\lim_{\gamma\to 0}\frac{\sigma_A^2}{\sigma_M^2} &= \left[\left\{1+\left(\frac{1-\theta}{\theta}\right)\left(\frac{\nu+e^{-\beta}}{\nu+1}\right)\right\}^{-1}+\left\{1+\left(\frac{\theta}{1-\theta}\right)\left(\frac{\nu+e^{\beta}}{\nu+1}\right)\right\}^{-1}\right]^{-1}+O(\rho) \\
&= 1+\frac{\nu\theta(1-\theta)}{(1+\nu)}\frac{(1-e^{\beta})^2}{\left\{(1-\theta+e^{\beta}\theta)^2+\nu e^{\beta}\right\}}+O(\rho) \\
&= \lambda_0+O(\rho),
\end{aligned}$$

where λ_0 is defined in (3.10) of the main text. Obviously, $\lambda_0 \geq 1$ and $\lambda_0 = 1$ if and only if $\beta = 0$.

Finally, we derive the asymptotic distribution of $\hat{\gamma}_{AC}$. The logarithm of the likelihood function (2.4) of the main text can be written as

$$\begin{aligned}
l_{AC} = \sum_{i=1}^{n} \big[&(\alpha+\beta x_i+\gamma g_i)d_i - \log(1+\exp(\alpha+\beta x_i+\gamma g_i)) \\
&+ x_i\log\theta + (1-x_i)\log(1-\theta) + g_i\log(\pi) + (1-g_i)\log(1-\pi)\big],
\end{aligned}$$

where θ is defined in (2.3) of the main text. It can be easily checked that the regularity conditions required for the asymptotic normality of $\hat{\gamma}_{AC}$ hold true (see Chapter 5, Van der Vaart, 2000). The Fisher information matrix is

$$I_{AC}(\mathbf{u}) = -E\frac{\partial^2 l_{AC}}{\partial\mathbf{u}\partial\mathbf{u}^T},$$

where $\mathbf{u} = (\alpha, \beta, \gamma, \pi)^T$. It is easy to derive that

$$I_{AC}(\mathbf{u}) = \begin{bmatrix} a & b & c & 0 \\ b & b & d & 0 \\ c & d & c & 0 \\ 0 & 0 & 0 & t \end{bmatrix} + g\frac{\partial\theta}{\partial\mathbf{u}}\frac{\partial\theta}{\partial\mathbf{u}^T} + h\frac{\partial^2\theta}{\partial\mathbf{u}\partial\mathbf{u}^T},$$

where

$$a = E\left(n_{+11}\frac{e^{\alpha+\beta+\gamma}}{(1+e^{\alpha+\beta+\gamma})^2} + n_{+10}\frac{e^{\alpha+\beta}}{(1+e^{\alpha+\beta})^2} + n_{+01}\frac{e^{\alpha+\gamma}}{(1+e^{\alpha+\gamma})^2} + n_{+00}\frac{e^{\alpha}}{(1+e^{\alpha})^2}\right),$$

$$b = E\left(n_{+11}\frac{e^{\alpha+\beta+\gamma}}{(1+e^{\alpha+\beta+\gamma})^2} + n_{+10}\frac{e^{\alpha+\beta}}{(1+e^{\alpha+\beta})^2}\right),$$

$$c = E\left(n_{+11}\frac{e^{\alpha+\beta+\gamma}}{(1+e^{\alpha+\beta+\gamma})^2} + n_{+01}\frac{e^{\alpha+\gamma}}{(1+e^{\alpha+\gamma})^2}\right),$$

$$d = E\left(n_{11+}\frac{e^{\alpha+\beta+\gamma}}{(1+e^{\alpha+\beta+\gamma})^2}\right),$$

$$t = E\left(\frac{n_{+11}+n_{+01}}{\pi^2} + \frac{n_{+10}+n_{+00}}{(1-\pi)^2}\right),$$

$$g = E\left(\frac{n_{+11}+n_{+10}}{\theta^2} + \frac{n_{+01}+n_{+00}}{(1-\theta)^2}\right),$$

$$h = E\left(\frac{n_{+01}+n_{+00}}{1-\theta} - \frac{n_{+11}+n_{+10}}{\theta}\right).$$

Since

$$E(n_{+ij}) = n_{1++}p_{1ij} + n_{0++}p_{0ij} = \frac{n}{1+\nu}\left(\nu p_{1ij} + p_{0ij}\right),$$

$$p_{1ij} = \mathrm{pr}(X=i, E=j \mid D=1) = (p_{ij}\theta^i(1-\theta)^{1-i}\pi^j(1-\pi)^{1-j})/f,$$

$$p_{0ij} = pr(X=i, E=j \mid D=0) = (q_{ij}\theta^i(1-\theta)^{1-i}\pi^j(1-\pi)^{1-j})/(1-f),$$

we have that

$$\lim_{\gamma\to 0}\lim_{f\to 0} a = \frac{e^{\alpha+\beta} n\theta}{1+\nu}\left(\frac{e^{\beta}\nu}{e^{\beta}\theta+1-\theta}+1\right)+\frac{e^{\alpha}n(1-\theta)}{1+\nu}\left(\frac{\nu}{e^{\beta}\theta+1-\theta}\right),$$

$$\lim_{\gamma\to 0}\lim_{f\to 0} b = \frac{e^{\alpha+\beta} n\theta}{1+\nu}\left(\frac{e^{\beta}\nu}{e^{\beta}\theta+1-\theta}+1\right),$$

$$\lim_{\gamma\to 0}\lim_{f\to 0} c = \frac{e^{\alpha+\beta} n\theta\pi}{1+\nu}\left(\frac{e^{\beta}\nu}{e^{\beta}\theta+1-\theta}+1\right)+\frac{e^{\alpha}n(1-\theta)\pi}{1+\nu}\left(\frac{\nu}{e^{\beta}\theta+1-\theta}\right),$$

$$\lim_{\gamma\to 0}\lim_{f\to 0} d = \frac{e^{\alpha+\beta} n\theta\pi}{1+\nu}\left(\frac{e^{\beta}\nu}{e^{\beta}\theta+1-\theta}+1\right),$$

$$\lim_{\gamma\to 0}\lim_{f\to 0} t = \frac{n}{\pi(1-\pi)},$$

$$\lim_{\gamma\to 0}\lim_{f\to 0} g = \frac{n}{\theta(1+\nu)}\left(\frac{e^{\beta}\nu}{e^{\beta}\theta+1-\theta}+1\right)+\frac{n}{(1-\theta)(1+\nu)}\left(\frac{\nu}{e^{\beta}\theta+1-\theta}+1\right),$$

and

$$\lim_{\gamma\to 0}\lim_{f\to 0} h = \frac{n\nu(1-e^{\beta})}{(1+\nu)(e^{\beta}\theta+1-\theta)}.$$

The standard likelihood theory gives that

$$n^{1/2}(\hat{\gamma}_{AC}-\gamma)\to N(0,\sigma^2_{AC}) \text{ in distribution,} \tag{S3.6}$$

where

$$\sigma^2_{AC} = n(I_{AC})^{-1}_{33}. \tag{S3.7}$$

After tedious symbolic algebra using the software Mathematica, we have that

$$\lim_{\gamma\to 0}\lim_{f\to 0} n(I_{AC})^{-1}_{33} = \sigma_0^2, \tag{S3.8}$$

where $\sigma_0^2 = (2+\nu+1/\nu)/\{\pi(1-\pi)\}$. That is,

$$\sigma^2_{AC}\to\sigma^2_0 \text{ as } f\to 0 \text{ and } \gamma\to 0. \tag{S3.9}$$

S4 Restatement of Pitman's asymptotic relative efficiency

We restate Pitman's asymptotic relative efficiency (Pitman, 1979; Serfling, 2009) below to facilitate our discussion in the main context.

Definition 1. Consider the problem of testing null hypothesis $H_0 : \gamma = 0$ against the alternative hypothesis $\gamma \neq 0$. For a sequence of test statistics indexed by sample size n, $T = \{T_n\}$, suppose that (i) there exist non-random variates $\mu_n(\gamma)$ and $\sigma_n(\gamma)$ such that $n^{1/2}(T_n - \mu_n(\gamma))/\sigma_n(\gamma)$ converges in distribution to the standard normal distribution as $n \to \infty$ under the contiguous alternative hypothesis $H_1 : \gamma = cn^{-1/2}$, (ii) $\mu_n(\gamma)$ has a continuous derivative $\mu_n'(\gamma)$ in a neighbourhood of 0, and (iii) $\sigma_n(\gamma)$ is continuous at 0. Then $n^{1/2}\sigma_n(0)/\mu_n'(0)$ converges to some constant as $n \to \infty$. Let κ_A and κ_B denote such constants corresponding to test statistic sequences T_A and T_B, respectively. Pitman's asymptotic relative efficiency of T_A to T_B is defined as $e_P(T_A, T_B) = (\kappa_B/\kappa_A)^2$.

S5 Proof of Theorem 1 and Corollary 3

Adopting the previous notations

$$\rho = e^{\alpha}, \ b_1 = e^{\beta}(1-\theta) + \theta, \ \text{and } b_2 = e^{\beta}/(e^{\beta}\theta - \theta + 1), \tag{S5.1}$$

then for δ defined in (S1.5) we have that

$$\begin{aligned}\lim_{\gamma\to 0}\frac{d(\gamma+\delta)}{d\gamma} &= \lim_{\gamma\to 0}\left[1+\frac{d}{d\gamma}\log\left\{1+\frac{e^{\alpha}(b_1-b_2)(1-e^{\gamma})}{(1+e^{\alpha+\gamma}b_1)(1+e^{\alpha}b_2)}\right\}\right]\\ &= 1-\frac{(b_1-b_2)\rho}{(1+b_1\rho)(1+b_2\rho)}\\ &= \frac{b_1b_2\rho^2+2b_2\rho+1}{b_1b_2\rho^2+(b_1+b_2)\rho+1}. \end{aligned}\tag{S5.2}$$

By (S5.2) and Lemma 2, Pitman's asymptotic relative efficiency of Mar to Adj is equal to

$$\begin{aligned}e_P(\hat{\gamma}_M,\hat{\gamma}_A) &= \left\{\lim_{\gamma\to 0}\left(\frac{d(\gamma+\delta)/d\gamma}{d\gamma/d\gamma}\right)\right\}^2\left\{\lim_{\gamma\to 0}\frac{\mathrm{var}(\hat{\gamma}_A)}{\mathrm{var}(\hat{\gamma}_M)}\right\}\\ &= \left\{\frac{b_1b_2\rho^2+2b_2\rho+1}{b_1b_2\rho^2+(b_1+b_2)\rho+1}\right\}^2\lambda.\end{aligned}$$

In the rare outcome situation ($f\to 0$ or equivalently $\rho\to 0$), applying Taylor's expansion to (S5.2), we have

$$\lim_{\gamma\to 0}\frac{d(\gamma+\delta)}{d\gamma} = 1-(b_1-b_2)\rho+(b_1^2-b_2^2)\rho^2+O(\rho^3). \tag{S5.3}$$

Consequently,

$$\left(\lim_{\gamma\to 0}\frac{d(\gamma+\delta)}{d\gamma}\right)^2 = 1-2(b_1-b_2)\rho+\{2(b_1^2-b_2^2)+(b_1-b_2)^2\}\rho^2+O(\rho^3). \tag{S5.4}$$

By (S5.4) and Corollary 2, when the outcome is rare, as indicated by a small ρ, Pitman's asymptotic relative efficiency of Mar to Adj is equal to

$$\begin{aligned} e_P(\hat{\gamma}_M, \hat{\gamma}_A) &= \left\{ \lim_{\gamma \to 0} \left(\frac{d(\gamma+\delta)/d\gamma}{d\gamma/d\gamma} \right) \right\}^2 \left\{ \lim_{\gamma \to 0} \frac{\mathrm{var}(\hat{\gamma}_A)}{\mathrm{var}(\hat{\gamma}_M)} \right\} \\ &= \left(1 - 2(b_1 - b_2)\rho + O(\rho^2)\right)(\lambda_0 + O(\rho)) \\ &= \lambda_0 + O(\rho), \end{aligned}$$

where λ_0 is defined in (3.10) of the main text.

S6 Proof of Theorem 2

We adopt the notations in the proof of Lemma 2:

$$\sigma_{AC}^2 = \lim_{n\to\infty} \mathrm{var}(n^{1/2}\hat{\gamma}_{AC}), \quad \sigma_M^2 = \lim_{n\to\infty} \mathrm{var}(n^{1/2}\hat{\gamma}_M).$$

The second-order Taylor expansion of σ_{AC}^2 with respect to f is

$$\sigma_{AC}^2 = \sigma_{AC}^2|_{f=0} + \frac{\partial}{\partial f}\sigma_{AC}^2\Big|_{f=0} \times f + \frac{1}{2}\frac{\partial^2}{\partial f^2}\sigma_{AC}^2\Big|_{f=0} \times f^2 + O(f^3),$$

so that

$$\begin{aligned} &e_P(\hat{\gamma}_M, \hat{\gamma}_{AC}) \\ =&\left(\lim_{\gamma \to 0} \frac{d(\gamma+\delta)/d\gamma}{d\gamma/d\gamma} \right)^2 \times \lim_{\gamma \to 0} \frac{\mathrm{var}(\hat{\gamma}_{AC})}{\mathrm{var}(\hat{\gamma}_M)} \\ =&\left\{ 1 - 2(b_1 - b_2)\rho + \left\{2(b_1^2 - b_2^2) + (b_1 - b_2)^2\right\}\rho^2 + O(\rho^3) \right\} \\ &\times \lim_{\gamma \to 0} \left(\frac{\sigma_{AC}^2|_{f=0}}{\sigma_M^2} + \frac{\frac{\partial}{\partial f}\sigma_{AC}^2|_{f=0}}{\sigma_M^2} \times f + \frac{1}{2}\frac{\frac{\partial^2}{\partial f^2}\sigma_{AC}^2|_{f=0}}{\sigma_M^2} \times f^2 + O(f^3) \right) \end{aligned} \tag{S6.1}$$

by (S5.4). Symbolic algebra with the software Mathematica gives

$$\frac{\partial}{\partial f}\sigma^2_{AC}\Big|_{f=0,\gamma=0} = \frac{2(2+\nu+1/\nu)(\theta-1)\theta\left(e^{\beta}-1\right)^2}{\left(\theta\left(e^{\beta}-1\right)+1\right)^2(\pi-1)\pi} \tag{S6.2}$$

and

$$\begin{aligned}
&\frac{\partial^2}{\partial f^2}\sigma^2_{AC}\Big|_{f=0,\gamma=0} \\
=& -\left[e^{\beta}\nu^2+\theta^2(e^{\beta}-1)^2(2\nu+1)-\theta(e^{\beta}-1\ )\{(e^{\beta}-3)\nu-2\}+5e^{\beta}\nu+\nu+1\right] \\
&\times\frac{2(\theta-1)\theta\left(e^{\beta}-1\right)^2(\nu+1)^2}{\left(\theta\left(e^{\beta}-1\right)+1\right)^4(\pi-1)\pi\nu^2}.
\end{aligned} \tag{S6.3}$$

If $\gamma = 0$, then the outcome prevalence can be expressed as

$$f|_{\gamma=0} = \frac{e^{\alpha+\beta}}{1+e^{\alpha+\beta}}\theta + \frac{e^{\alpha}}{1+e^{\alpha}}(1-\theta),$$

so that

$$f|_{\gamma=0} = (e^{\beta}\theta-\theta+1)\rho - (e^{2\beta}\theta-\theta+1)\rho^2 + O(\rho^3). \tag{S6.4}$$

It follows from (S3.2) and (S3.9) that

$$\sigma^2_M|_{\gamma=0} = \sigma^2_0 \text{ and } \sigma^2_{AC}|_{f=0,\gamma=0} = \sigma^2_0. \tag{S6.5}$$

Furthermore, from (S6.2)-(S6.4), we have

$$\begin{aligned}
&\frac{\frac{\partial}{\partial f}\sigma^2_{AC}|_{f=0,\gamma=0}}{\sigma^2_0}\times f \\
&= \frac{2(1-\theta)\theta\left(e^{\beta}-1\right)^2}{\theta\left(e^{\beta}-1\right)+1}\rho + \frac{2(1-\theta)\theta(e^{\beta}-1)^2(\theta-1-e^{2\beta}\theta)}{(\theta(e^{\beta}-1)+1)^2}\rho^2 + O(\rho^3) \\
&= 2(b_1-b_2)\rho - \frac{2(1-\theta)\theta(e^{\beta}-1)^2(1+(e^{2\beta}-1)\theta)}{(\theta(e^{\beta}-1)+1)^2}\rho^2 + O(\rho^3)
\end{aligned} \tag{S6.6}$$

and

$$\frac{1}{2}\frac{\frac{\partial^2}{\partial f^2}\sigma^2_{AC}|_{f=0}}{\sigma_0^2}\times f^2$$
$$=\big[(e^\beta\nu^2+\theta^2(e^\beta-1)^2(2\nu+1)-\theta(e^\beta-1)\{(e^\beta-3)\nu-2\}+5e^\beta\nu+\nu+1\big]$$
$$\times\frac{(\theta-1)\theta(e^\beta-1)^2}{\{\theta(e^\beta-1)+1\}^2\nu}\rho^2+O(\rho^3). \tag{S6.7}$$

By equations (S6.1) and (S6.5)-(S6.7), we have

$$e_P(\hat\gamma_M,\hat\gamma_{AC})=1+\tau\rho^2+O(\rho^3),$$

where

$$\tau=-\frac{(1-\theta)\theta(e^\beta-1)^2\{(1+1/\nu)[(\theta(e^\beta-1)+1)^2+e^\beta\nu]+2(1+(e^{2\beta}-1)\theta)\}}{\{\theta\,(e^\beta-1)+1\}^2}. \tag{S6.8}$$

Obviously, $\tau\le 0$ and the equality holds if and only if $\beta=0$.

S7 Robustness of AdjCon with respect to prevalence specification

S7.1 Notations and preliminary results

Let $\boldsymbol{s}=(\beta,\gamma,\theta,\pi)^\top$ denote unknown model parameters. Denote by f_0 the true outcome prevalence, which is incorrectly specified as f_1 in AdjCon. In this section, we show that AdjCon is robust with respect to the misspecification of the outcome prevalence. Let B be the domain of $(\beta,\gamma,\theta,\pi)$:

$B = \{(\beta, \gamma, \theta, \pi) \mid \beta$ and γ are bounded away from infinity, θ and π are bounded away from zero and one$\}$.

Assume that $f_0, f_1 \in (0, 1-\epsilon]$ for some give $\epsilon > 0$, which is easily hold in practice. Assume that $\boldsymbol{s}_f^* = (\beta^*, \gamma^*, \theta^*, \pi^*) \in B$ for any $f \in (0, 1-\epsilon]$. The log-likelihood is

$$l_f(\boldsymbol{s}) = (\alpha + \beta X + \gamma E)D - \log(1 + \exp(\alpha + \beta X + \gamma E)) +$$
$$X \log\theta + (1-X)\log(1-\theta) + E\log(\pi) + (1-E)\log(1-\pi),$$

subject to the prevalence constraint

$$f = \sum_{i=0}^{1}\sum_{j=0}^{1} p(D=1 \mid X=i, E=j)p(X=i)p(E=j). \tag{S7.1}$$

We will show that both $\boldsymbol{s}_{f_1}^*$ and the corresponding asymptotic covariance matrix $\Sigma_{f_1}(\boldsymbol{s}_{f_1}^*)$ are Lipschitz continuous with respect to f_1, that is

$$\|\boldsymbol{s}_{f_1}^* - \boldsymbol{s}_{f_0}^*\| = C_1|f_1 - f_0|$$

and

$$\|\Sigma_{f_1}(\boldsymbol{s}_{f_1}^*) - \Sigma_{f_0}(\boldsymbol{s}_{f_0}^*)\| = C_2|f_1 - f_0|, \tag{S7.2}$$

where

$$\boldsymbol{s}_{f_0}^* = \arg\max_{\boldsymbol{s}} E_{f_0}\{l_{f_0}(\boldsymbol{s})\}, \quad \boldsymbol{s}_{f_1}^* = \arg\max_{\boldsymbol{s}} E_{f_0}\{l_{f_1}(\boldsymbol{s})\}, \tag{S7.3}$$

and C_1, C_2 are independent of f_0 and f_1.

We define the following quantities that will be used in the proof of Lemma 1. Let

$$M_1(\boldsymbol{s}) = e^{\beta+\gamma}\theta\pi + e^{\beta}\theta(1-\pi) + e^{\gamma}(1-\theta)\pi + (1-\theta)(1-\pi),$$

then

$$0 < m_1 = \min_{\boldsymbol{s}\in B}\{e^{\beta+\gamma}, e^{\beta}, e^{\gamma}, 1\} \leq M_1(\boldsymbol{s}) \leq \max_{\boldsymbol{s}\in B}\{e^{\beta+\gamma}, e^{\beta}, e^{\gamma}, 1\} = M_1.$$

Let

$$M_2(\boldsymbol{s}) = e^{-\beta-\gamma}\theta\pi + e^{-\beta}\theta(1-\pi) + e^{-\gamma}(1-\theta)\pi + (1-\theta)(1-\pi),$$

then

$$0 < m_2 = \min_{\boldsymbol{s}\in B}\{e^{-\beta-\gamma}, e^{-\beta}, e^{-\gamma}, 1\} \leq M_2(\boldsymbol{s}) \leq \max_{\boldsymbol{s}\in B}\{e^{-\beta-\gamma}, e^{-\beta}, e^{-\gamma}, 1\} = M_2.$$

The following lemma presents a decomposition of the intercept parameter α under the prevalence constraint.

Lemma 1. *Assume $f \in (0, 1-\epsilon]$ for some $\epsilon > 0$, β and γ are bounded away from infinity and θ and π are bounded away from zero and one. Denote $\boldsymbol{s} = (\beta, \gamma, \theta, \pi)$. The intercept α, as a function of f and $\boldsymbol{s}$ due to constraint*

$$f = \sum_{i=0}^{1}\sum_{j=0}^{1} pr(D = 1 \mid X = i, E = j)pr(X = i)pr(E = j), \qquad \text{(S7.4)}$$

can be decomposed into two parts:

$$\alpha(f, \boldsymbol{s}) = \alpha_1(f) + \alpha_2(f, \boldsymbol{s}),$$

where $\alpha_2(f, \boldsymbol{s})$ is Lipschitz continuous with respect to f.

Proof. Denote $\rho(f, \boldsymbol{s}) = \exp(\alpha(f, \boldsymbol{s}))$ and rewrite the constraint (S7.4):

$$\begin{aligned}
f =& F(\alpha, \boldsymbol{s}) = \sum_{i=0}^{1}\sum_{j=0}^{1} p(D=1|X=i, E=j)p(X=i)p(E=j) \\
=& \frac{\exp(\alpha+\beta+\gamma)\theta\pi}{1+\exp(\alpha+\beta+\gamma)} + \frac{\exp(\alpha+\beta)\theta(1-\pi)}{1+\exp(\alpha+\beta)} + \frac{\exp(\alpha+\gamma)(1-\theta)\pi}{1+\exp(\alpha+\gamma)} \\
&+ \frac{\exp(\alpha)(1-\theta)(1-\pi)}{1+\exp(\alpha)} \\
=& \frac{\rho}{1+\rho}\left(\frac{(1+\rho)e^{\beta+\gamma}\theta\pi}{1+\rho e^{\beta+\gamma}} + \frac{(1+\rho)e^{\beta}\theta(1-\pi)}{1+\rho e^{\beta}} + \frac{(1+\rho)e^{\gamma}(1-\theta)\pi}{1+\rho e^{\gamma}} + (1-\theta)(1-\pi)\right) \\
=& \frac{\rho}{1+\rho}C'(\rho, \boldsymbol{s}), \qquad \text{(S7.5)}
\end{aligned}$$

where as ρ ranges from 0 to ∞, $C'(\rho, \boldsymbol{s})$ ranges from $M_1(\boldsymbol{s})$ to 1. Similarly,

$$\begin{aligned}
1-f =& \frac{1}{1+\rho}\left[\frac{(1+\rho)\theta\pi}{1+\rho e^{\beta+\gamma}} + \frac{(1+\rho)\theta(1-\pi)}{1+\rho e^{\beta}} + \frac{(1+\rho)(1-\theta)\pi}{1+\rho e^{\gamma}} + (1-\theta)(1-\pi)\right] \\
=& \frac{1}{1+\rho}C''(\rho, \boldsymbol{s}), \qquad \text{(S7.6)}
\end{aligned}$$

where as ρ ranges from 0 to ∞, $C''(\rho, \boldsymbol{s})$ ranges from 1 to $M_2(\boldsymbol{s})$. Combining (S7.5) and (S7.6), we have

$$\begin{aligned}
\alpha(f, \boldsymbol{s}) &= \{\log f - \log(1-f)\} + \{\log C''(\rho(f, \boldsymbol{s}), \boldsymbol{s}) - \log C'(\rho(f, \boldsymbol{s}), \boldsymbol{s})\} \\
&= \alpha_1(f) + \alpha_2(f, \boldsymbol{s}).
\end{aligned}$$

In what follows, we show that

$$\alpha_2(f, \boldsymbol{s}) = \log C''(\rho(f, \boldsymbol{s}), \boldsymbol{s}) - \log C'(\rho(f, \boldsymbol{s}), \boldsymbol{s}) \qquad \text{(S7.7)}$$

is Lipschitz continuous with respect to f. First,

$$\begin{aligned}
&\left|\frac{\partial \log C''(\rho,\boldsymbol{s})}{\partial f}\right| = \left|\frac{1}{C''(\rho,\boldsymbol{s})}\frac{\partial C''(\rho,\boldsymbol{s})}{\partial \rho}\frac{\partial \rho}{\partial f}\right| \\
=&\left|\frac{1}{C''(\rho,\boldsymbol{s})}\frac{\frac{\theta\pi(1-e^{\beta+\gamma})}{(1+\rho e^{\beta+\gamma})^2}+\frac{\theta(1-\pi)(1-e^{\beta})}{(1+\rho e^{\beta})^2}+\frac{(1-\theta)\pi(1-e^{\gamma})}{(1+\rho e^{\gamma})^2}}{\frac{e^{\beta+\gamma}\theta\pi}{(1+\rho e^{\beta+\gamma})^2}+\frac{e^{\beta}\theta(1-\pi)}{(1+\rho e^{\beta})^2}+\frac{e^{\gamma}(1-\theta)\pi}{(1+\rho e^{\gamma})^2}+\frac{(1-\theta)(1-\pi)}{(1+\rho)^2}}\right| \\
=&\frac{1}{C''(\rho,\boldsymbol{s})}\frac{1}{C'''(\rho,\boldsymbol{s})}\left|\frac{(1+\rho)^2\theta\pi}{(1+\rho e^{\beta+\gamma})^2}+\frac{(1+\rho)^2\theta(1-\pi)}{(1+\rho e^{\beta})^2}+\frac{(1+\rho)^2(1-\theta)\pi}{(1+\rho e^{\gamma})^2}\right. \\
&\left.+(1-\theta)(1-\pi)-C'''(\rho,\boldsymbol{s})\right| \\
\leq&\frac{1}{m_2}\frac{1}{\min\{m_1,m_2\}}(M_2^2+M_1+M_2),
\end{aligned} \tag{S7.8}$$

where

$$\begin{aligned}
C'''(\rho,\boldsymbol{s}) = \Bigg\{&\frac{(1+\rho)^2e^{\beta+\gamma}\theta\pi}{(1+\rho e^{\beta+\gamma})^2}+\frac{(1+\rho)^2e^{\beta}\theta(1-\pi)}{(1+\rho e^{\beta})^2}+\frac{(1+\rho)^2e^{\gamma}(1-\theta)\pi}{(1+\rho e^{\gamma})^2} \\
&+(1-\theta)(1-\pi)\Bigg\}.
\end{aligned}$$

As ρ ranges from 0 to ∞, $C'''(\rho,\boldsymbol{s})$ ranges from $M_1(\boldsymbol{s})$ to $M_2(\boldsymbol{s})$. Second,

$$\begin{aligned}
&\left|\frac{\partial \log C'(\rho,\boldsymbol{s})}{\partial f}\right| = \left|\frac{1}{C'(\rho,\boldsymbol{s})}\frac{\partial C'(\rho,\boldsymbol{s})}{\partial \rho}\frac{\partial \rho}{\partial f}\right| \\
=&\left|\frac{1}{C'(\rho,\boldsymbol{s})}\frac{\frac{e^{\beta+\gamma}\theta\pi(1-e^{\beta+\gamma})}{(1+\rho e^{\beta+\gamma})^2}+\frac{e^{\beta}\theta(1-\pi)(1-e^{\beta})}{(1+\rho e^{\beta})^2}+\frac{e^{\gamma}(1-\theta)\pi(1-e^{\gamma})}{(1+\rho e^{\gamma})^2}}{\frac{e^{\beta+\gamma}\theta\pi}{(1+\rho e^{\beta+\gamma})^2}+\frac{e^{\beta}\theta(1-\pi)}{(1+\rho e^{\beta})^2}+\frac{e^{\gamma}(1-\theta)\pi}{(1+\rho e^{\gamma})^2}+\frac{(1-\theta)(1-\pi)}{(1+\rho)^2}}\right| \\
=&\frac{1}{C'(\rho,\boldsymbol{s})}\frac{1}{C'''(\rho,\boldsymbol{s})}\left|C'''(\rho,\boldsymbol{s})-\left[\frac{e^{2\beta+2\gamma}(1+\rho)^2\theta\pi}{(1+\rho e^{\beta+\gamma})^2}+\frac{e^{2\beta}(1+\rho)^2\theta(1-\pi)}{(1+\rho e^{\beta})^2}\right.\right. \\
&\left.\left.+\frac{e^{2\gamma}(1+\rho)^2(1-\theta)\pi}{(1+\rho e^{\gamma})^2}+(1-\theta)(1-\pi)\right]\right| \\
\leq&\frac{1}{m_1}\frac{1}{\min\{m_1,m_2\}}(M_1^2+M_1+M_2).
\end{aligned} \tag{S7.9}$$

It follows from (S7.7)-(S7.9) that $\alpha_2(f, \boldsymbol{s})$ is Lipschitz continuous with respect to $f \in (0, 1-\epsilon]$ for some $\epsilon > 0$ and any $\boldsymbol{s} \in B$. Denote the Lipschitz constant by

$$L_C = \frac{1}{\min\{m_1, m_2\}}\left(\frac{1}{m_2}(M_2^2 + M_1 + M_2) + \frac{1}{m_1}(M_1^2 + M_1 + M_2)\right). \quad \text{(S7.10)}$$

□

S7.2 Proof of Theorem 3: part I (Lipschitz continuity of $\boldsymbol{s}^*_{f_1}$)

Before proving $\|\boldsymbol{s}^*_{f_1} - \boldsymbol{s}^*_{f_0}\| \le C_1|f_1 - f_0|$, we first prove

$$E_{f_0}[l_{f_0}(\boldsymbol{s}^*_{f_0}) - l_{f_0}(\boldsymbol{s}^*_{f_1})] \le C|f_1 - f_0|. \quad \text{(S7.11)}$$

Since $E_{f_0}[l_{f_1}(\boldsymbol{s}^*_{f_0})] \le E_{f_0}[l_{f_1}(\boldsymbol{s}^*_{f_1})]$, we have

$$E_{f_0}[l_{f_0}(\boldsymbol{s}^*_{f_0}) + (l_{f_1}(\boldsymbol{s}^*_{f_0}) - l_{f_0}(\boldsymbol{s}^*_{f_0}))] \le E_{f_0}[l_{f_0}(\boldsymbol{s}^*_{f_1}) + (l_{f_1}(\boldsymbol{s}^*_{f_1}) - l_{f_0}(\boldsymbol{s}^*_{f_1}))]. \quad \text{(S7.12)}$$

Thus,

$$0 \le E_{f_0}[l_{f_0}(\boldsymbol{s}^*_{f_0}) - l_{f_0}(\boldsymbol{s}^*_{f_1})] \le E_{f_0}[(l_{f_1}(\boldsymbol{s}^*_{f_1}) - l_{f_0}(\boldsymbol{s}^*_{f_1}))] - E_{f_0}[(l_{f_1}(\boldsymbol{s}^*_{f_0}) - l_{f_0}(\boldsymbol{s}^*_{f_0}))],$$

where the first equality holds according to the definition of (S7.3). We need only to prove that the right-hand side of the above inequality is Lipschitz continuous with respect to f.

For any $\boldsymbol{s} \in B$, we have

$$
\begin{aligned}
E_{f_0}[l_{f_1}(\boldsymbol{s}) - l_{f_0}(\boldsymbol{s})] &= E_{f_0}\left[\frac{\partial l_f(\boldsymbol{s})}{\partial f}\bigg|_{f=f^*}\right](f_1 - f_0) \\
&= E_{f_0}\left[(\alpha_1 - \alpha_0)D - \log\frac{1+\exp(\alpha_1 + \beta X + \gamma E)}{1+\exp(\alpha_0 + \beta X + \gamma E)}\right] \\
&= \nu/(1+\nu)(\alpha_1 - \alpha_0) - E_{f_0}\left[\log\frac{1+\exp(\alpha_1 + \beta X + \gamma E)}{1+\exp(\alpha_0 + \beta X + \gamma E)}\right],
\end{aligned}
$$

so that

$$
\begin{aligned}
& E_{f_0}[(l_{f_1}(\boldsymbol{s}^*_{f_1}) - l_{f_0}(\boldsymbol{s}^*_{f_1}))] - E_{f_0}[(l_{f_1}(\boldsymbol{s}^*_{f_0}) - l_{f_0}(\boldsymbol{s}^*_{f_0}))] \\
=& E_{f_0}\left[\nu/(1+\nu)(\alpha(f_1, \boldsymbol{s}^*_{f_1}) - \alpha(f_0, \boldsymbol{s}^*_{f_1})) - \log\frac{1+\exp(\alpha(f_1, \boldsymbol{s}^*_{f_1}) + \beta X + \gamma E)}{1+\exp(\alpha(f_0, \boldsymbol{s}^*_{f_1}) + \beta X + \gamma E)}\right] \\
& - E_{f_0}\left[\nu/(1+\nu)(\alpha(f_1, \boldsymbol{s}^*_{f_0}) - \alpha(f_0, \boldsymbol{s}^*_{f_0})) - \log\frac{1+\exp(\alpha(f_1, \boldsymbol{s}^*_{f_0}) + \beta X + \gamma E)}{1+\exp(\alpha(f_0, \boldsymbol{s}^*_{f_0}) + \beta X + \gamma E)}\right] \\
=& \frac{\nu}{1+\nu}\big([\alpha(f_1, \boldsymbol{s}^*_{f_1}) - \alpha(f_0, \boldsymbol{s}^*_{f_1})] - [\alpha(f_1, \boldsymbol{s}^*_{f_0}) - \alpha(f_0, \boldsymbol{s}^*_{f_0})]\big) \\
& - \left\{E_{f_0}\left[\frac{\rho^*_1\exp(\beta^*_1 X + \gamma^*_1 E)}{1+\rho^*_1\exp(\beta^*_1 X + \gamma^*_1 E)}\right]\frac{\partial\alpha(f, \boldsymbol{s}^*_{f_1})}{\partial f}\bigg|_{f=f^*_1}\right. \\
& \left. - E_{f_0}\left[\frac{\rho^*_0\exp(\beta^*_0 X + \gamma^*_0 E)}{1+\rho^*_0\exp(\beta^*_0 X + \gamma^*_0 E)}\right]\frac{\partial\alpha(f, \boldsymbol{s}^*_{f_0})}{\partial f}\bigg|_{f=f^*_0}\right\}(f_1 - f_0). \qquad \text{(S7.13)}
\end{aligned}
$$

When $0 < f_1, f_0 < 1-\epsilon, \boldsymbol{s}^*_{f_1}, \boldsymbol{s}^*_{f_0} \in B$, in what follows, we show both of the two terms in the righthand side of (S7.13) are Lipschitz continuous with respect to f. First, according to Lemma 1,

$$
\alpha(f, \boldsymbol{s}) = \alpha_1(f) + \alpha_2(f, \boldsymbol{s}). \qquad \text{(S7.14)}
$$

For the first term in (S7.13),

$$
\begin{aligned}
&[\alpha(f_1, \boldsymbol{s}^*_{f_1}) - \alpha(f_0, \boldsymbol{s}^*_{f_1})] - [\alpha(f_1, \boldsymbol{s}^*_{f_0}) - \alpha(f_0, \boldsymbol{s}^*_{f_0})] \\
=&[\alpha(f_1, \boldsymbol{s}^*_{f_1}) - \alpha(f_1, \boldsymbol{s}^*_{f_0})] - [\alpha(f_0, \boldsymbol{s}^*_{f_1}) - \alpha(f_0, \boldsymbol{s}^*_{f_0})] \\
=&[\alpha_2(f_1, \boldsymbol{s}^*_{f_1}) - \alpha_2(f_0, \boldsymbol{s}^*_{f_1})] - [\alpha_2(f_1, \boldsymbol{s}^*_{f_0}) - \alpha_2(f_0, \boldsymbol{s}^*_{f_0})] \\
\leq& 2L_C|f_1 - f_0|
\end{aligned}
$$

holds because $\alpha_2(f, \boldsymbol{s})$ is Lipschitz continuous with respect to f. Next, we show the second term is also Lipschitz continuous with respect to f. Note that

$$
\begin{aligned}
\frac{\partial \alpha(f, \boldsymbol{s})}{\partial f} = \Bigg\{ \frac{\rho}{(1+\rho)^2} \Bigg[& \frac{(1+\rho)^2 e^{\beta+\gamma}\theta\pi}{(1+\rho e^{\beta+\gamma})^2} + \frac{(1+\rho)^2 e^{\beta}\theta(1-\pi)}{(1+\rho e^{\beta})^2} + \frac{(1+\rho)^2 e^{\gamma}(1-\theta)\pi}{(1+\rho e^{\gamma})^2} \\
& + (1-\theta)(1-\pi) \Bigg] \Bigg\}^{-1} \\
= \frac{(1+\rho)^2}{\rho C'''(\rho, \boldsymbol{s})} = \frac{1+\rho}{\rho} & \frac{1+\rho}{C'''(\rho, \boldsymbol{s})} = \frac{1+\rho}{\rho} \frac{C''(\rho, \boldsymbol{s})}{(1-f)C'''(\rho, \boldsymbol{s})}.
\end{aligned}
\tag{S7.15}
$$

According to Equation (S7.15), the second term is

$$
\begin{aligned}
&\Bigg\{ E_{f_0} \left[\frac{(1+\rho_1^*)\exp(\beta_1^* X + \gamma_1^* E)}{1+\rho_1^* \exp(\beta_1^* X + \gamma_1^* E)} \right] \frac{C''(\rho_1^*, \boldsymbol{s}^*_{f_1})}{(1-f_1^*)C'''(\rho_1^*, \boldsymbol{s}^*_{f_1})} \\
&\quad - E_{f_0} \left[\frac{(1+\rho_0^*)\exp(\beta_0^* X + \gamma_0^* E)}{1+\rho_0^* \exp(\beta_0^* X + \gamma_0^* E)} \right] \frac{C''(\rho_0^*, \boldsymbol{s}^*_{f_0})}{(1-f_0^*)C'''(\rho_0^*, \boldsymbol{s}^*_{f_0})} \Bigg\} (f_1 - f_0) \\
\leq& 2L_1 L_\epsilon L_3 |f_1 - f_0|,
\end{aligned}
$$

where

$$
L_1 = \max_{\boldsymbol{s}^* \in B} \left\{ 1, E_{f_0} \exp(\beta^* X + \gamma^* E) \right\} \leq \max_{\boldsymbol{s}^* \in B} \{ 1, \exp(\beta^* + \gamma^*), \exp(\beta^*), \exp(\gamma^*) \} = M_1,
$$

$$L_\epsilon = \max\left\{\frac{1}{1-f_1}, \frac{1}{1-f_0}\right\} \le \frac{1}{\epsilon}, \quad \text{if } f_1, f_0 \le 1-\epsilon,$$

$$L_3 = \max_{\rho^*>0, \boldsymbol{s}^*\in B} \frac{C''(\rho^*, \boldsymbol{s}^*)}{C'''(\rho^*, \boldsymbol{s}^*)} = \frac{M_2}{\min\{m_1, m_2\}}.$$

So we have

$$E_{f_0}\{l_{f_0}(\boldsymbol{s}^*_{f_0}) - l_{f_0}(\boldsymbol{s}^*_{f_1})\} \le C|f_1 - f_0|,$$

where $C = 2\nu L_C/(1+\nu) + 2L_1 L_\epsilon L_3$.

Taylor's expansion of $l_{f_0}(\boldsymbol{s}^*_{f_1})$ at $\boldsymbol{s}^*_{f_0}$ gives that

$$l_{f_0}(\boldsymbol{s}^*_{f_1}) = l_{f_0}(\boldsymbol{s}^*_{f_0}) + (\boldsymbol{s}^*_{f_1} - \boldsymbol{s}^*_{f_0})^\top \frac{\partial l_{f_0}(\boldsymbol{s})}{\partial \boldsymbol{s}}\bigg|_{\boldsymbol{s}=\boldsymbol{s}^*_{f_0}} + (\boldsymbol{s}^*_{f_1} - \boldsymbol{s}^*_{f_0})^\top \frac{\partial^2 l_{f_0}(\boldsymbol{s})}{\partial \boldsymbol{s}\partial \boldsymbol{s}^\top}\bigg|_{\boldsymbol{s}=\boldsymbol{s}'} (\boldsymbol{s}^*_{f_1} - \boldsymbol{s}^*_{f_0}), \tag{S7.16}$$

where $\boldsymbol{s}'$ lies in between $\boldsymbol{s}^*_{f_0}$ and $\boldsymbol{s}^*_{f_1}$. Consequently,

$$\begin{aligned} E_{f_0}\{l_{f_0}(\boldsymbol{s}^*_{f_0}) - l_{f_0}(\boldsymbol{s}^*_{f_1})\} &= (\boldsymbol{s}^*_{f_1} - \boldsymbol{s}^*_{f_0})^\top \left\{ - E_{f_0} \frac{\partial^2 l_{f_0}(\boldsymbol{s})}{\partial \boldsymbol{s}\partial \boldsymbol{s}^\top}\bigg|_{\boldsymbol{s}=\boldsymbol{s}'} \right\} (\boldsymbol{s}^*_{f_1} - \boldsymbol{s}^*_{f_0}) \\ &\le C|f_1 - f_0|. \end{aligned} \tag{S7.17}$$

It can be easily show that the matrix $-E_{f_0}\{\partial^2 l_{f_0}(\boldsymbol{s})/(\partial \boldsymbol{s}\partial \boldsymbol{s}^\top)|_{\boldsymbol{s}=\boldsymbol{s}'}\}$ is positive definite. Let the corresponding smallest eigenvalue be $\lambda_{min} > 0$, then (S7.17) implies

$$\|\boldsymbol{s}^*_{f_1} - \boldsymbol{s}^*_{f_0}\| \le \frac{C}{\lambda_{min}}|f_1 - f_0|. \tag{S7.18}$$

Finally, let $C_1 = C/\lambda_{\min}$ which is given in Theorem 3 of the paper.

S7.3 Proof of Theorem 3: part II (Lipschitz continuity of $\Sigma_{f_1}(\boldsymbol{s}^*_{f_1})$)

We prove the asymptotic covariance matrix $\Sigma_{f_1}(\boldsymbol{s}^*_{f_1})$ is Lipschitz continuous with respect to f_1. Let $\hat{\boldsymbol{s}}_{f_1}$ maximize the log-likelihood function

$$l_{n,f_1} = \sum_{i=1}^{n} \Big[(\alpha_1 + \beta x_i + \gamma g_i)d_i - \log(1 + \exp(\alpha_1 + \beta x_i + \gamma g_i)) + x_i \log\theta + (1 - x_i)\log(1-\theta) + g_i \log(\pi) + (1 - g_i)\log(1-\pi)\Big]$$

with the outcome prevalence being specified to be f_1.

According to White (1982), the maximum likelihood estimator $\hat{\boldsymbol{s}}_{f_1}$ is consistent for $\boldsymbol{s}^*_{f_1}$ and asymptotically normal:

$$\sqrt{n}(\hat{\boldsymbol{s}}_{f_1} - \boldsymbol{s}^*_{f_1}) \to N(0, \Sigma_{f_1}(\boldsymbol{s}^*_{f_1})), \tag{S7.19}$$

where

$$\Sigma_{f_1}(\boldsymbol{s}^*_{f_1}) = A^{-1}(f_1, \boldsymbol{s}^*_{f_1})B(f_1, \boldsymbol{s}^*_{f_1})A^{-1}(f_1, \boldsymbol{s}^*_{f_1}) \tag{S7.20}$$

with

$$A(f, \boldsymbol{s}) = \frac{1}{n}E_{f_0}\left\{\frac{\partial^2 l_{n,f}(\boldsymbol{s})}{\partial \boldsymbol{s}\partial \boldsymbol{s}^\top}\right\} \text{ and } B(f, \boldsymbol{s}) = \frac{1}{n}E_{f_0}\left\{\frac{\partial l_{n,f}(\boldsymbol{s})}{\partial \boldsymbol{s}}\frac{\partial l_{n,f}(\boldsymbol{s})}{\partial \boldsymbol{s}^\top}\right\}. \tag{S7.21}$$

Assume that $A(f, \boldsymbol{s})$ and $B(f, \boldsymbol{s})$ have good condition numbers among all $f \in (0, 1-\epsilon]$ and $\boldsymbol{s} \in B$. Specifically, $\|A(f, \boldsymbol{s})\| \le \Lambda_A, \|A^{-1}(f, \boldsymbol{s})\| \le \lambda_A$ and $\|B(f, \boldsymbol{s})\| \le \Lambda_B, \|B^{-1}(f, \boldsymbol{s})\| \le \lambda_B$. Similarly, we have

$$\Sigma_{f_0}(\boldsymbol{s}^*_{f_0}) = A^{-1}(f_0, \boldsymbol{s}^*_{f_0})B(f_0, \boldsymbol{s}^*_{f_0})A^{-1}(f_0, \boldsymbol{s}^*_{f_0}). \tag{S7.22}$$

Note that

$$-A(f_0, \boldsymbol{s}^*_{f_0}) = B(f_0, \boldsymbol{s}^*_{f_0}). \tag{S7.23}$$

In order to show $\|\Sigma_{f_1}(\boldsymbol{s}^*_{f_1}) - \Sigma_{f_0}(\boldsymbol{s}^*_{f_0})\| \leq C_2|f_1 - f_0|$, we only need to show

$$\|A(f_1, \boldsymbol{s}^*_{f_1}) - A(f_0, \boldsymbol{s}^*_{f_0})\| \leq C_A|f_1 - f_0|, \tag{S7.24}$$

and

$$\|B(f_1, \boldsymbol{s}^*_{f_1}) - B(f_0, \boldsymbol{s}^*_{f_0})\| \leq C_B|f_1 - f_0|. \tag{S7.25}$$

In fact, according to the Woodbury matrix identity

$$(A - B)^{-1} = A^{-1} + A^{-1}B(A - B)^{-1}$$

that

$$\begin{aligned} A^{-1}(f_0, \boldsymbol{s}^*_{f_0}) &= (A(f_1, \boldsymbol{s}^*_{f_1}) - [A(f_1, \boldsymbol{s}^*_{f_1}) - A(f_0, \boldsymbol{s}^*_{f_0})])^{-1} \\ &= A^{-1}(f_1, \boldsymbol{s}^*_{f_1}) + A^{-1}(f_1, \boldsymbol{s}^*_{f_1})[A(f_1, \boldsymbol{s}^*_{f_1}) - A(f_0, \boldsymbol{s}^*_{f_0})]A^{-1}(f_0, \boldsymbol{s}^*_{f_0}) \end{aligned} \tag{S7.26}$$

Consequently,

$$\begin{aligned} \|A^{-1}(f_1, \boldsymbol{s}^*_{f_1}) - A^{-1}(f_0, \boldsymbol{s}^*_{f_0})\| &\leq \|A^{-1}(f_1, \boldsymbol{s}^*_{f_1})\|\|A(f_1, \boldsymbol{s}^*_{f_1}) - A(f_0, \boldsymbol{s}^*_{f_0})\|\|A^{-1}(f_0, \boldsymbol{s}^*_{f_0})\| \\ &\leq \lambda_A^2 C_A|f_1 - f_0|. \end{aligned} \tag{S7.27}$$

By (S7.20), (S7.22), (S7.23), (S7.24) and (S7.27), we have

$$
\begin{aligned}
&\|\Sigma_{f_1}(\boldsymbol{s}^*_{f_1}) - \Sigma_{f_0}(\boldsymbol{s}^*_{f_0})\| \\
=&\|A^{-1}(f_1, \boldsymbol{s}^*_{f_1})B(f_1, \boldsymbol{s}^*_{f_1})A^{-1}(f_1, \boldsymbol{s}^*_{f_1}) - A^{-1}(f_0, \boldsymbol{s}^*_{f_0})B(f_0, \boldsymbol{s}^*_{f_0})A^{-1}(f_0, \boldsymbol{s}^*_{f_0})\| \\
=&\|A^{-1}(f_1, \boldsymbol{s}^*_{f_1})B(f_1, \boldsymbol{s}^*_{f_1})A^{-1}(f_1, \boldsymbol{s}^*_{f_1}) - A^{-1}(f_1, \boldsymbol{s}^*_{f_1})B(f_0, \boldsymbol{s}^*_{f_0})A^{-1}(f_1, \boldsymbol{s}^*_{f_1}) \\
&+ A^{-1}(f_1, \boldsymbol{s}^*_{f_1})B(f_0, \boldsymbol{s}^*_{f_0})A^{-1}(f_1, \boldsymbol{s}^*_{f_1}) - A^{-1}(f_0, \boldsymbol{s}^*_{f_0})B(f_0, \boldsymbol{s}^*_{f_0})A^{-1}(f_0, \boldsymbol{s}^*_{f_0})\| \\
\leq&\|A^{-1}(f_1, \boldsymbol{s}^*_{f_1})\|\|B(f_1, \boldsymbol{s}^*_{f_1}) - B(f_0, \boldsymbol{s}^*_{f_0})\|\|A^{-1}(f_1, \boldsymbol{s}^*_{f_1})\| \\
&+ \|A^{-1}(f_1, \boldsymbol{s}^*_{f_1}) - A^{-1}(f_0, \boldsymbol{s}^*_{f_0})\|\|B(f_0, \boldsymbol{s}^*_{f_0})\|\|A^{-1}(f_1, \boldsymbol{s}^*_{f_1}) + A^{-1}(f_0, \boldsymbol{s}^*_{f_0})\| \\
\leq&(\lambda_A^2 C_B + 2\lambda_A^3 C_A \Lambda_B)|f_1 - f_0|.
\end{aligned}
\tag{S7.28}
$$

Here we let $C_2 = \lambda_A^2 C_B + 2\lambda_A^3 C_A \Lambda_B$ which is given in Theorem 3 of the main text.

Now we prove Equations (S7.24) and (S7.25). Given the prevalence constraint (S7.5), i.e., $f = F(\alpha, \boldsymbol{s})$, we have

$$
\frac{\partial \alpha(f, \boldsymbol{s})}{\partial \boldsymbol{s}} = -\frac{\partial F/\partial \boldsymbol{s}}{\partial F/\partial \alpha}(f, \boldsymbol{s}).
$$

It can be verified that when $f \in (0, 1-\epsilon]$, $\boldsymbol{s} \in B$, $\partial \alpha(f, \boldsymbol{s})/\partial \boldsymbol{s}$ is bounded

Lipschitz continuous with respect to f, and the derivative

$$\frac{\partial l_f(\boldsymbol{s})}{\partial \boldsymbol{s}} = \begin{bmatrix} (D - \frac{\exp(\alpha+\beta X+\gamma E)}{1+\exp(\alpha+\beta X+\gamma E)})(X + \frac{\partial \alpha}{\partial \beta}) \\ (D - \frac{\exp(\alpha+\beta X+\gamma E)}{1+\exp(\alpha+\beta X+\gamma E)})(E + \frac{\partial \alpha}{\partial \gamma}) \\ \frac{X}{\theta} - \frac{1-X}{1-\theta} + (D - \frac{\exp(\alpha+\beta X+\gamma E)}{1+\exp(\alpha+\beta X+\gamma E)})\frac{\partial \alpha}{\partial \theta} \\ \frac{E}{\pi} - \frac{1-E}{1-\pi} + (D - \frac{\exp(\alpha+\beta X+\gamma E)}{1+\exp(\alpha+\beta X+\gamma E)})\frac{\partial \alpha}{\partial \pi} \end{bmatrix}$$

is also bounded Lipschitz continuous with respect to f, since $\exp(\alpha(f,\boldsymbol{s}))/(1+\exp(\alpha(f,\boldsymbol{s})))$ is bounded Lipschitz continuous with respect to f. By the fact that the product of two bounded Lipschitz continuous functions is also bounded Lipschitz continuous, we have $\{\partial l_f(\boldsymbol{s})/\partial \boldsymbol{s}\}\{\partial l_f(\boldsymbol{s})/\partial \boldsymbol{s}^\top\}$ is Lipschitz continuous with respect to f (assume the Lipschitz constant L_{Bf}). Moreover, $\{\partial l_f(\boldsymbol{s})/\partial \boldsymbol{s}\}\{\partial l_f(\boldsymbol{s})/\partial \boldsymbol{s}^\top\}$ is a continuously differentiable function with respect to $\boldsymbol{s}$ in the compact set B, so $B(f,\boldsymbol{s})$ is Lipschitz continuous with respect to $\boldsymbol{s}$ (assume the Lipschitz constant L_{Bs}). Using $\|\boldsymbol{s}^*_{f_1} - \boldsymbol{s}^*_{f_0}\| \le C_1|f_1 - f_0|$ proved in Section S7.2, we have that Equation (S7.25) holds:

$$\begin{aligned}\|B(f_1,\boldsymbol{s}^*_{f_1}) - B(f_0,\boldsymbol{s}^*_{f_0})\| &\le \|B(f_1,\boldsymbol{s}^*_{f_1}) - B(f_0,\boldsymbol{s}^*_{f_1})\| + \|B(f_0,\boldsymbol{s}^*_{f_1}) - B(f_0,\boldsymbol{s}^*_{f_0})\| \\ &\le (L_{Bf} + L_{Bs}C_1)|f_1 - f_0|.\end{aligned}$$

Let $C_B = L_{Bf} + L_{Bs}C_1$ which is defined in (S7.25).

Similarly,

$$\frac{\partial^2 \alpha(f,\boldsymbol{s})}{\partial \boldsymbol{s}\partial \boldsymbol{s}^\top} = -\frac{\frac{\partial F}{\partial \alpha}(\frac{\partial \alpha}{\partial \boldsymbol{s}})^2 + 2\frac{\partial^2 F}{\partial \alpha \partial \boldsymbol{s}}\frac{\partial \alpha}{\partial \boldsymbol{s}} + \frac{\partial^2 F}{\partial \boldsymbol{s}\partial \boldsymbol{s}^\top}}{\frac{\partial F}{\partial \alpha}}$$

is also bounded Lipschitz continuous with respect to f. Denote

$$l_1(f,\boldsymbol{s}) = \frac{\partial l(\alpha,\boldsymbol{s})}{\partial \alpha}, l_2(f,\boldsymbol{s}) = \frac{\partial l(\alpha,\boldsymbol{s})}{\partial \boldsymbol{s}},$$

$$l_{11}(f,\boldsymbol{s}) = \frac{\partial^2 l(\alpha,\boldsymbol{s})}{\partial \alpha \partial \alpha},\ l_{12}(f,\boldsymbol{s}) = \frac{\partial^2 l(\alpha,\boldsymbol{s})}{\partial \alpha \partial \boldsymbol{s}},\ l_{22}(f,\boldsymbol{s}) = \frac{\partial^2 l(\alpha,\boldsymbol{s})}{\partial \boldsymbol{s} \partial \boldsymbol{s}}.$$

Since $l_f(\boldsymbol{s}) = l(\alpha(f,\boldsymbol{s}),\boldsymbol{s})$, we have

$$\begin{aligned}
\frac{\partial^2 l_f(\boldsymbol{s})}{\partial \boldsymbol{s} \partial \boldsymbol{s}^\top} &= \frac{\partial}{\partial \boldsymbol{s}}\left[l_1(f,\boldsymbol{s})\frac{\partial \alpha}{\partial \boldsymbol{s}} + l_2(f,\boldsymbol{s})\right] \\
&= \left[l_{11}(f,\boldsymbol{s})\frac{\partial \alpha}{\partial \boldsymbol{s}} + l_{12}(f,\boldsymbol{s})\right]\frac{\partial \alpha}{\partial \boldsymbol{s}} + l_1(f,\boldsymbol{s})\frac{\partial^2 \alpha}{\partial \boldsymbol{s}\partial \boldsymbol{s}^\top} + l_{12}(f,\boldsymbol{s})\frac{\partial \alpha}{\partial \boldsymbol{s}} + l_{22}(f,\boldsymbol{s}) \\
&= l_{11}(f,\boldsymbol{s})\frac{\partial \alpha}{\partial \boldsymbol{s}}\frac{\partial \alpha}{\partial \boldsymbol{s}^\top} + 2l_{12}(f,\boldsymbol{s})\frac{\partial \alpha}{\partial \boldsymbol{s}^\top} + l_1(f,\boldsymbol{s})\frac{\partial^2 \alpha}{\partial \boldsymbol{s}\partial \boldsymbol{s}^\top} + l_{22}(f,\boldsymbol{s})
\end{aligned}$$

is bounded Lipschitz continuous since each of the terms in the right hand side of the above equation is a bounded Lipschitz continuous function with respect to f. Also $\partial^2 l_f(\boldsymbol{s})/(\partial \boldsymbol{s}\partial \boldsymbol{s}^\top)$ is continuously differentiable with respect to $\boldsymbol{s}$ in the compact region B, thus $A(f,\boldsymbol{s})$ is also Lipschitz continuous with respect to $\boldsymbol{s}$. So we have that (S7.24) holds:

$$\begin{aligned}
\|A(f_1,\boldsymbol{s}^*_{f_1}) - A(f_0,\boldsymbol{s}^*_{f_0})\| &\le \|A(f_1,\boldsymbol{s}^*_{f_1}) - A(f_0,\boldsymbol{s}^*_{f_1})\| + \|A(f_0,\boldsymbol{s}^*_{f_1}) - A(f_0,\boldsymbol{s}^*_{f_0})\| \\
&\le C_A|f_1 - f_0|.
\end{aligned}$$

Finally, following from (S7.24) and (S7.25), we have (S7.28) holds.

Table S1: Type-I error rate/power with possibly misspecified f.

	$\gamma = 0$			$\gamma = 0.075$		
f_0	Mar	Adj	AdjCon	Mar	Adj	AdjCon
0.01	0.050	0.050	0.050	0.755	0.733	0.754
0.05	0.052	0.051	0.051	0.748	0.730	0.747
0.10	0.049	0.049	0.049	0.737	0.730	0.737
0.20	0.050	0.050	0.050	0.728	0.730	0.731

In any of the four scenarios, the prevalence f is specified to be 0.05 in AdjCon.

S8 Simulation study for robustness of AdjCon

In this study, we assess the robustness of AdjCon against disease prevalence misspecification. Case-control data are generated in a manner analogous to Section 4 of the main text, with $\gamma = 0$ or 0.075 and $f_0 = 0.01$, 0.05, 0.10, and 0.2. When applying AdjCon, we specify the disease prevalence to be 0.05, so that the disease prevalence is correctly specified when $f_0 = 0.05$ and misspecified otherwise. We also include Mar and Adj for the purpose of comparison. The resulting type-I error rates ($\gamma = 0$) and powers ($\gamma = 0.075$) are presented in Figure S1, which are obtained based on 50,000 replications. As demonstrated in Figure S1(A), AdjCon maintains well-controlled type-I error rates, even in scenarios where the disease prevalence is significantly misestimated. Furthermore, AdjCon is at lest comparative in terms of powers against the two alternative methods in the presence of prevalence misspecfication (Figure S1(B)). These empirical results coincide with the theoretical insights discussed in Section 3.3 of the main text.

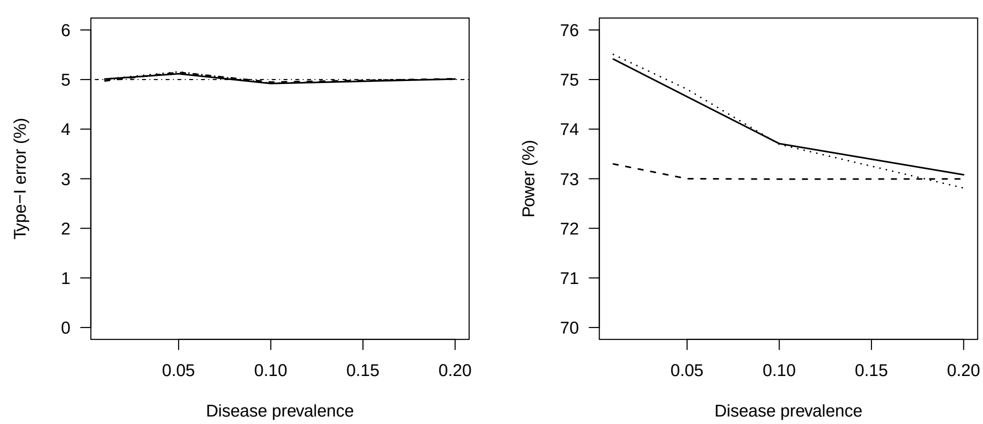

Figure S1: (A) Type-I error rates of MAR (dotted line), ADJ (dashed line), and ADJCON (solid line) for testing exposure-disease association ($H_0 : \gamma = 0$) with $\gamma = 0$, $\beta = 1$, $\theta = \pi = 0.5$, $n_0 = n_1 = 10000$; (B) Powers of MAR (dotted line), ADJ (dashed line), and ADJCON (solid line) for testing exposure-disease association ($H_0 : \gamma = 0$) with $\gamma = 0.075$, $\beta = 1$, $\theta = \pi = 0.5$, $n_0 = n_1 = 10000$.

S9 Additional results for the HDL-C data analysis

Table S2: P-values for SNP vs. BMI association tests

SNP	P-values	SNP	P-values	SNP	P-values	SNP	P-values
rs2144300	5.70E-01	rs4846914	5.70E-01	rs3779788	9.02E-01	rs255	9.99E-01
rs256	9.41E-01	rs263	7.42E-01	rs264	8.70E-01	rs271	9.76E-01
rs301	3.45E-01	rs328	6.93E-01	rs331	6.94E-01	rs12679834	7.95E-01
rs3208305	4.51E-01	rs3735964	8.03E-01	rs13702	4.74E-01	rs3916027	7.37E-01
rs2197089	2.40E-01	rs1340510	3.60E-01	rs3890182	5.47E-01	rs2275544	8.67E-01
rs1883025	5.61E-01	rs7120118	4.14E-01	rs102275	2.22E-01	rs2338104	5.46E-01
rs11635491	1.96E-02	rs1800588	1.55E-01	rs2070895	1.40E-01	rs8034802	5.26E-02
rs8033940	6.15E-02	rs261332	1.07E-01	rs588136	9.56E-02	rs261341	3.89E-02
rs261338	2.25E-01	rs13306677	1.59E-01	rs6499861	1.69E-02	rs6499863	6.29E-03
rs12708967	1.55E-01	rs3764261	1.51E-02	rs12720918	1.58E-01	rs17231506	1.51E-02
rs4783961	3.05E-02	rs1800775	4.41E-03	rs711752	1.43E-02	rs708272	1.43E-02
rs1864163	8.31E-03	rs7203984	1.38E-02	rs11508026	2.26E-02	rs12720922	3.84E-02
rs9939224	2.97E-02	rs11076174	3.89E-01	rs1532625	1.08E-02	rs1532624	6.00E-03
rs11076175	2.03E-02	rs7499892	1.11E-02	rs11076176	7.25E-01	rs289714	7.69E-01
rs5880	4.13E-01	rs1800777	6.67E-02	rs2292318	2.37E-01	rs255052	1.32E-01
rs1943981	1.08E-02	rs2156552	1.10E-02	rs2075650	2.53E-01	rs6073952	7.99E-01

Grayed are p-values smaller than 0.05.

Table S3: P-values for SNP vs. HDL-C association tests

SNP	Mar	Adj	AdjCon	SNP	Mar	Adj	AdjCon
rs2144300	1.07E-01	1.65E-01	1.37E-01	rs4846914	1.07E-01	1.65E-01	1.37E-01
rs3779788	**6.10E-04**	**3.17E-04**	**2.33E-04**	rs255	7.16E-03	2.80E-03	2.56E-03
rs256	1.73E-03	**6.43E-04**	**4.83E-04**	rs263	**1.24E-04**	**3.44E-05**	**2.33E-05**
rs264	1.64E-03	**6.18E-04**	**5.27E-04**	rs271	5.87E-03	2.43E-03	2.00E-03
rs301	1.23E-03	3.64E-03	2.41E-03	rs328	8.71E-04	8.94E-04	**5.02E-04**
rs331	1.79E-02	2.63E-02	2.19E-02	rs12679834	1.36E-03	1.16E-03	**7.05E-04**
rs3208305	**2.69E-04**	**3.98E-04**	**2.51E-04**	rs3735964	4.97E-03	3.77E-03	2.69E-03
rs13702	**2.83E-04**	**3.46E-04**	**2.28E-04**	rs3916027	1.21E-02	1.57E-02	1.29E-02
rs2197089	2.07E-01	2.52E-02	3.33E-02	rs1340510	6.77E-02	3.60E-02	3.40E-02
rs3890182	4.94E-02	2.35E-02	2.77E-02	rs2275544	2.86E-02	2.09E-02	2.36E-02
rs1883025	1.87E-02	2.99E-02	2.79E-02	rs7120118	9.75E-01	5.41E-01	6.20E-01
rs102275	1.24E-01	3.06E-01	2.53E-01	rs2338104	7.26E-01	5.26E-01	5.56E-01
rs1800588	1.77E-03	5.39E-03	3.98E-03	rs2070895	1.30E-03	5.10E-03	3.61E-03
rs8034802	9.94E-03	9.08E-02	5.64E-02	rs8033940	7.62E-03	6.47E-02	4.06E-02
rs261332	1.08E-03	7.75E-03	4.72E-03	rs588136	3.87E-03	2.68E-02	1.68E-02
rs261338	2.76E-02	6.99E-02	5.40E-02	rs13306677	8.00E-01	2.16E-01	2.92E-01
rs12708967	1.12E-02	7.53E-03	6.56E-03	rs12720918	2.14E-02	2.41E-02	2.17E-02
rs11076174	**2.77E-06**	**4.36E-06**	**3.07E-06**	rs11076176	**8.81E-09**	**9.49E-10**	**8.12E-10**
rs289714	**8.22E-09**	**4.95E-10**	**4.73E-10**	rs5880	8.25E-03	1.23E-02	1.43E-02
rs1800777	7.69E-03	4.03E-02	3.17E-02	rs2292318	8.96E-01	4.28E-01	5.10E-01
rs255052	9.03E-01	5.66E-01	6.58E-01	rs2075650	7.80E-02	2.93E-02	3.24E-02
rs6073952	5.37E-01	4.53E-01	4.46E-01				

Results are present only for those SNPs not significantly associated with BMI at level 0.05. Bolded are significant results at level 0.05 after Bonferroni correction (p-value $< 0.05/64$).

S10 Simulation results for the probit link

In this section, we evaluate the three considered methods Mar, Adj, and AdjCon through simulations with the probit link function. The data generation process is the same as that in Section 4, except that the logit link function is replaced with the probit link function:

$$\text{pr}(D = 1 \mid X = i, E = j) = \Phi(\alpha + \beta i + \gamma j), \tag{S10.29}$$

where Φ is the cumulative function of the standard normal distribution. The parameters are set as $\beta = 1.0, \gamma = 0$ or 0.04, and $f = 0.01, 0.05, 0.1, 0.2,$ 0.25, or 0.3. A population of size 10^7 is generated for each parameter combination, and a sample of $n_{1++} = 10,000$ cases and $n_{0++} = 10,000$ controls are sampled from diseased and non-diseased individuals, respectively. Wald test statistics for AdjCon, Mar, and Adj are calculated for each generated dataset. Type-I error rates ($\gamma = 0$) and powers ($\gamma = 0.04$) under the nominal level 0.05 are obtained based on 100,000 simulation replicates. Figure S2 presents the corresponding type-I error rates and powers for the three methods.

As shown in Figure S2(A), all methods maintain well controlled type-I error rates around the nominal level 0.05. Furthermore, Figure S2(B) shows power trends of the three methods similar to those under the logit link function (Figure 1(C) and Figure 2(B)). Specifically, AdjCon is uniformly more powerful than Mar and Adj across various disease prevalences, while Mar is more powerful than Adj for small f and vice versa for large f.

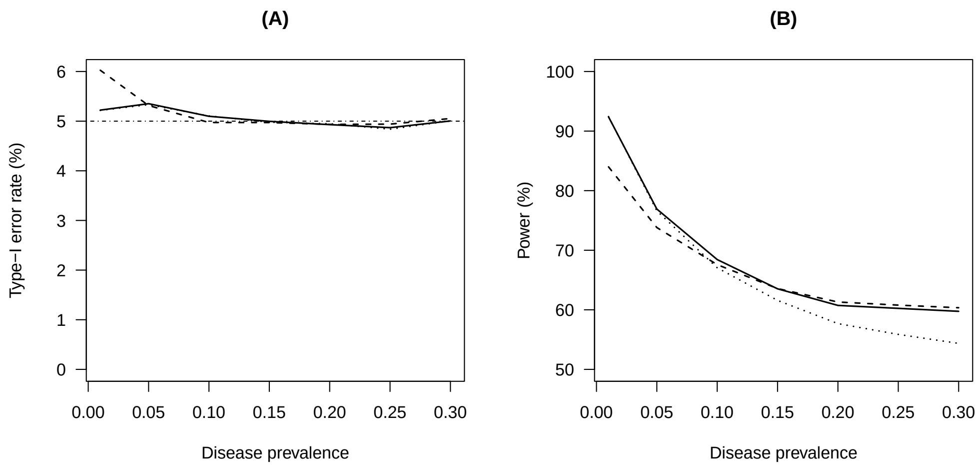

Figure S2: (A) Type-I error rates of Mar (dotted line), Adj (dashed line), and AdjCon (solid line) for testing exposure-disease association under the probit link ($H_0 : \gamma = 0$) with $\gamma = 0$, $\beta = 1$, $\theta = \pi = 0.5$, $n_0 = n_1 = 10000$; (B) Powers of Mar (dotted line), Adj (dashed line), and AdjCon (solid line) for testing exposure-disease association ($H_0 : \gamma = 0$) with $\gamma = 0.04$, $\beta = 1$, $\theta = \pi = 0.5$, $n_0 = n_1 = 10000$.

References

Gart, J. J. (1962). On the combination of relative risks. Biometrics 18(4), 601–610.

Pitman, E. J. G. (1979). Some Basic Theory of Statistical Inference. London: Chapman & Hall.

REFERENCES

Serfling, R. J. (2009). Approximation Theorems of Mathematical Statistics, Volume 162. John Wiley & Sons.

Van der Vaart, A. W. (2000). Asymptotic Statistics, Volume 3. Cambridge University Press.

White, H. (1982). Maximum likelihood estimation of misspecified models. Econometrica 50(1), 1–25.