Submitted to the Annals of Applied Statistics

APPENDIXES AND SUPPLEMENTARY TABLES: STATISTICAL METHODS FOR ANALYSIS OF COMBINED CATEGORICAL BIOMARKER DATA FROM MULTIPLE STUDIES

By Chao Cheng[†], Bernard Rosner[‡] and Molin Wang[‡,*]

Yale University[†] *and Harvard University*[‡]

[*]Corresponding author

APPENDIX SECTION

Appendix A: Estimating Equations in the Exact Calibration Method. In $\boldsymbol{\Psi}(\boldsymbol{\pi})$, $\Psi_{\boldsymbol{\theta}}$ and $\Psi_{\boldsymbol{\sigma}^2}$ are defined in (2.9). Additionally, $\Psi_{\boldsymbol{\beta}} = [\Psi_{\boldsymbol{\beta_0}}^T, \Psi_{\boldsymbol{\beta}_x}^T, \Psi_{\boldsymbol{\beta}_z}^T]^T$ takes the forms

$$\Psi_{\boldsymbol{\beta_0}} = \left(\sum_{k=1}^{N_1} \frac{\partial}{\partial\beta_{01}} \log \tilde{L}_{1k}, \dots, \sum_{k=1}^{N_M} \frac{\partial}{\partial\beta_{0M}} \log \tilde{L}_{Mk}\right)^T,$$
$$\Psi_{\boldsymbol{\beta}_x} = \left(\sum_{j=1}^{M}\sum_{k=1}^{N_j} \frac{\partial}{\partial\beta_{x,2}} \log \tilde{L}_{jk}, \dots, \sum_{j=1}^{M}\sum_{k=1}^{N_j} \frac{\partial}{\partial\beta_{x,G}} \log \tilde{L}_{jk}\right)^T,$$
$$\Psi_{\boldsymbol{\beta}_z} = \left(\sum_{j=1}^{M}\sum_{k=1}^{N_j} \frac{\partial}{\partial\beta_{z_1}} \log \tilde{L}_{jk}, \dots, \sum_{j=1}^{M}\sum_{k=1}^{N_j} \frac{\partial}{\partial\beta_{z_P}} \log \tilde{L}_{jk}\right)^T,$$

where

$$\frac{\partial}{\partial\beta_{0j}} \log \tilde{L}_{jk} = \sum_{l=2}^{G} \hat{p}_{jk,l} \frac{Y_{jk}\exp\{Y_{jk}(\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})\}+(Y_{jk}-1)\exp\{(Y_{jk}+1)(\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})\}}{\tilde{L}_{jk}[1+\exp(\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})]^2},$$
$$\frac{\partial}{\partial\beta_{x,l}} \log \tilde{L}_{jk} = \hat{p}_{jk,l} \frac{Y_{jk}\exp\{Y_{jk}(\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})\}+(Y_{jk}-1)\exp\{(Y_{jk}+1)(\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})\}}{\tilde{L}_{jk}[1+\exp(\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})]^2},$$
$$\frac{\partial}{\partial\beta_{z_p}} \log \tilde{L}_{jk} = \sum_{l=2}^{G} \hat{p}_{jk,l} \frac{Z_{jk,p}Y_{jk}\exp\{Y_{jk}(\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})\}+Z_{jk,p}(Y_{jk}-1)\exp\{(Y_{jk}+1)(\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})\}}{\tilde{L}_{jk}[1+\exp(\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})]^2}.$$

Appendix B: Approximate Likelihood Derivation in the Cut-off Calibration Method. From equation (2.5), the likelihood contribution from the k^{th} individual in the j^{th} study is

$$L_{jk} = \sum_{l=2}^{G} \frac{\exp\{Y_{jk}(\beta_{0j} + \beta_{x,l} + \boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})\}}{1+\exp(\beta_{0j} + \beta_{x,l} + \boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})} p_{jk,l}, \tag{A.1}$$

where $p_{jk,l} = P(g_{l-1} \le X_{jk} < g_l|\boldsymbol{H}_{jk}, \boldsymbol{W}_{jk})$. In this section, we prove that under certain conditions, the likelihood contribution above can be well approximated by

$$L_{jk}^{(c)} = \frac{\exp\{Y_{jk}(\beta_{0j} + \sum_{l=2}^{G}\mathbb{I}(g_{l-1} \le \mu_{jk} < g_l)\beta_{x,l} + \boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})\}}{1+\exp(\beta_{0j} + \sum_{l=2}^{G}\mathbb{I}(g_{l-1} \le \mu_{jk} < g_l)\beta_{x,l} + \boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})}, \tag{A.2}$$

which is the likelihood contribution used by the cut-off method. The first condition involves small local laboratory measurement errors, i.e $\sigma_d^2 \approx 0$ $(d = 1, \dots, M)$. For brevity, we suppress some subscripts. Under this condition and based on (2.7) and (2.8), we have $\mathrm{var}(X_{jk}|\boldsymbol{H}_{jk}, \boldsymbol{W}_{jk}) \approx 0$, i.e $s_{jk}^2 \approx 0$, for individuals in both calibration and non-calibration subset. In (2.11), when s_{jk} is small enough,

$$p_{jk,l} \approx \begin{cases} 1 & \text{if } g_{l-1} \le \mu_{jk} < g_l \\ 0 & \text{else.} \end{cases} \tag{A.3}$$

It follows that $L_{jk} \approx L_{jk}^{(c)}$. The second condition assumes small exposure effect; that is, all $\beta_{x,l} \approx 0$. Under this condition, both L_{jk} and $L_{jk}^{(c)}$ are approximately equal to $\frac{\exp\{Y_{jk}(\beta_{0j}+\boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})\}}{1+\exp(\beta_{0j}+\boldsymbol{\beta}_z^T\boldsymbol{Z}_{jk})}$. Thus $L_{jk} \approx L_{jk}^{(c)}$. Therefore, the approximated likelihood (A.2) is close to (A.1) if all local laboratory measurement errors are small and/or the biomarker effect is weak.

Appendix C: Estimating Equations in the Cut-off Calibration Method. Let $\boldsymbol{\beta}^T\boldsymbol{X}^* = \beta_{0j} + \sum_{l=2}^{G} \mathbb{I}(g_{l-1} \leq \mu_{jk} < g_l)\beta_{x,l} + \boldsymbol{\beta}_z^T \boldsymbol{Z}_{jk}$ and $\boldsymbol{\pi} = [\boldsymbol{\theta}^T, {\boldsymbol{\sigma}^2}^T, \boldsymbol{\beta}^T]^T$. We define the estimating equations $\boldsymbol{\Psi}^{(c)}(\boldsymbol{\pi}) = [\Psi_{\boldsymbol{\theta}}, \Psi_{\boldsymbol{\sigma}^2}, \Psi_{\boldsymbol{\beta}}^{(c)}]$, where $\Psi_{\boldsymbol{\theta}}$ and $\Psi_{\boldsymbol{\sigma}^2}$ are given in (2.9), and $\Psi_{\boldsymbol{\beta}}^{(c)} = [\Psi_{\boldsymbol{\beta_0}}^{(c)T}, \Psi_{\boldsymbol{\beta}_x}^{(c)T}, \Psi_{\boldsymbol{\beta}_z}^{(c)T}]^T$ takes the forms

$$\begin{aligned}
\Psi_{\boldsymbol{\beta_0}}^{(c)} &= \Big(\textstyle\sum_{k=1}^{N_1} \frac{\partial}{\partial \beta_{01}} \log \tilde{L}_{1k}^{(c)}, \ldots, \sum_{k=1}^{N_M} \frac{\partial}{\partial \beta_{0M}} \log \tilde{L}_{Mk}^{(c)} \Big)^T, \\
\Psi_{\boldsymbol{\beta}_x}^{(c)} &= \Big(\textstyle\sum_{j=1}^{M}\sum_{k=1}^{N_j} \frac{\partial}{\partial \beta_{x,2}} \log \tilde{L}_{jk}^{(c)}, \ldots, \sum_{j=1}^{M}\sum_{k=1}^{N_j} \frac{\partial}{\partial \beta_{x,G}} \log \tilde{L}_{jk}^{(c)} \Big)^T, \\
\Psi_{\boldsymbol{\beta}_z}^{(c)} &= \Big(\textstyle\sum_{j=1}^{M}\sum_{k=1}^{N_j} \frac{\partial}{\partial \beta_{z_1}} \log \tilde{L}_{jk}^{(c)}, \ldots, \sum_{j=1}^{M}\sum_{k=1}^{N_j} \frac{\partial}{\partial \beta_{z_P}} \log \tilde{L}_{jk}^{(c)} \Big)^T,
\end{aligned}$$

with

$$\tilde{L}_{jk}^{(c)} = \frac{\exp\{Y_{jk}(\beta_{0j} + \sum_{l=2}^{G} \mathbb{I}(g_{l-1} \leq \hat{\mu}_{jk} < g_l)\beta_{x,l} + \boldsymbol{\beta}_z^T \boldsymbol{Z}_{jk})\}}{1 + \exp(\beta_{0j} + \sum_{l=2}^{G} \mathbb{I}(g_{l-1} \leq \hat{\mu}_{jk} < g_l)\beta_{x,l} + \boldsymbol{\beta}_z^T \boldsymbol{Z}_{jk})},$$

$$\begin{aligned}
\frac{\partial}{\partial \beta_{0j}} \log \tilde{L}_{jk}^{(c)} &= \tfrac{Y_{jk}\exp\{Y_{jk}(\boldsymbol{\beta}^T\boldsymbol{X}^*)\} + (Y_{jk}-1)\exp\{(Y_{jk}+1)(\boldsymbol{\beta}^T\boldsymbol{X}^*)\}}{\tilde{L}_{jk}^{(c)}[1+\exp(\boldsymbol{\beta}^T\boldsymbol{X}^*)]^2} \\
\frac{\partial}{\partial \beta_{x,l}} \log \tilde{L}_{jk}^{(c)} &= \mathbb{I}(g_{l-1} \leq \hat{\mu}_{jk} < g_l)\tfrac{Y_{jk}\exp\{Y_{jk}(\boldsymbol{\beta}^T\boldsymbol{X}^*)\} + (Y_{jk}-1)\exp\{(Y_{jk}+1)(\boldsymbol{\beta}^T\boldsymbol{X}^*)\}}{\tilde{L}_{jk}^{(c)}[1+\exp(\boldsymbol{\beta}^T\boldsymbol{X}^*)]^2} \\
\frac{\partial}{\partial \beta_{z_p}} \log \tilde{L}_{jk}^{(c)} &= Z_{jk,p}\tfrac{Y_{jk}\exp\{Y_{jk}(\boldsymbol{\beta}^T\boldsymbol{X}^*)\} + (Y_{jk}-1)\exp\{(Y_{jk}+1)(\boldsymbol{\beta}^T\boldsymbol{X}^*)\}}{\tilde{L}_{jk}^{(c)}[1+\exp(\boldsymbol{\beta}^T\boldsymbol{X}^*)]^2}
\end{aligned}$$

We use the standard sandwich variance method to estimate the variance of $\boldsymbol{\pi}$. Specifically, we have

$$\widehat{\text{var}}(\hat{\boldsymbol{\pi}}) = \hat{\boldsymbol{Q}}^{-1}\hat{\boldsymbol{U}}(\hat{\boldsymbol{Q}}^{-1})^T,$$

where $\hat{\boldsymbol{Q}} = \sum_{j=1}^{M}\sum_{k=1}^{N_j}\Big[\frac{d\boldsymbol{\psi}_{\boldsymbol{\pi},jk}^{(c)}}{d\boldsymbol{\pi}}|_{\boldsymbol{\pi}=\hat{\boldsymbol{\pi}}}\Big]$, $\hat{\boldsymbol{U}} = \sum_{j=1}^{M}\sum_{k=1}^{N_j}\Big[\boldsymbol{\psi}_{\hat{\boldsymbol{\pi}},jk}^{(c)} {\boldsymbol{\psi}_{\hat{\boldsymbol{\pi}},jk}}^{(c)T}\Big]$, and $\boldsymbol{\psi}_{\boldsymbol{\pi},jk}^{(c)}$ is the term in $\boldsymbol{\Psi}^{(c)}(\boldsymbol{\pi})$ corresponding to the k^{th} individual in the j^{th} study.

Appendix D: Calibration Parameters under Controls Only Calibration. In this section, we show that under either of the following two conditions, we have $P(X_{jk}|\boldsymbol{H}_{jk}, \boldsymbol{W}_{jk}, Y_{jk} = 0) \approx P(X_{jk}|\boldsymbol{H}_{jk}, \boldsymbol{W}_{jk})$, $\alpha_{0j,co} \approx \alpha_{0j}$ and $\boldsymbol{\tau}_{co} \approx \boldsymbol{\tau}$. The first condition is small or null exposure effect, and the second condition is rare disease prevalence. In the beginning, we can show that $E(H_{jk,d}|\boldsymbol{W}_{jk}, Y_{jk} = 0)$ can be written as (some subscripts are suppressed for brevity)

$$\begin{aligned}
E(H|\boldsymbol{W}, Y = 0) &= \int H f(H|\boldsymbol{W}, Y = 0)dh \\
&= \int H \frac{f(H, Y = 0|\boldsymbol{W})}{\int f(H, Y = 0|\boldsymbol{W})dh} dh \\
&= \int H \frac{f(Y = 0|H, \boldsymbol{W}) f(H|\boldsymbol{W})}{\int f(Y = 0|H, \boldsymbol{W}) f(H|\boldsymbol{W})dh} dh,
\end{aligned}$$

and $f(X|\boldsymbol{H},\boldsymbol{W},Y=0)$ can be shown as

$$\begin{aligned}
f(X|\boldsymbol{H},\boldsymbol{W},Y=0) &= \frac{f(X,\boldsymbol{H},\boldsymbol{W},Y=0)}{f(\boldsymbol{H},\boldsymbol{W},Y=0)} \\
&= \frac{f(Y=0|X,\boldsymbol{H},\boldsymbol{W})f(X,\boldsymbol{H},\boldsymbol{W})}{\int f(Y=0|X,\boldsymbol{H},\boldsymbol{W})f(X,\boldsymbol{H},\boldsymbol{W})dx} \\
&= \frac{f(Y=0|X,\boldsymbol{Z}^*)f(X|\boldsymbol{H},\boldsymbol{W})}{\int f(Y=0|X,\boldsymbol{Z}^*)f(X|\boldsymbol{H},\boldsymbol{W})dx},
\end{aligned}$$

where $\boldsymbol{Z}^*=\boldsymbol{Z}\cap\boldsymbol{W}$ denotes the common variates of $\boldsymbol{Z}$ and $\boldsymbol{W}$. Notice that $E(H|\boldsymbol{W},Y=0)\approx E(H|\boldsymbol{W})$ indicate $\alpha_{0j,co}\approx\alpha_{0j}$ and $\boldsymbol{\tau}_{co}\approx\boldsymbol{\tau}$. We only need to prove $E(H|\boldsymbol{W},Y=0)\approx E(H|\boldsymbol{W})$ and $f(X|\boldsymbol{H},\boldsymbol{W},Y=0)\approx f(X|\boldsymbol{H},\boldsymbol{W})$. Now, we prove that under either conditions.

Condition I: Small exposure effect. If $\boldsymbol{\beta}_x\approx\mathbf{0}$, Y is nearly independent with X, i.e $f(Y=0|X,\boldsymbol{W})\approx f(Y=0|\boldsymbol{W})$, which implies

$$\begin{aligned}
f(Y=0|H,\boldsymbol{W}) &= \int f(Y=0|H,X,\boldsymbol{W})f(X|H,\boldsymbol{W})dx \\
&= \int f(Y=0|X,\boldsymbol{W})f(X|H,\boldsymbol{W})dx \\
&\approx \int f(Y=0|\boldsymbol{W})f(X|H,\boldsymbol{W})dx \\
&= f(Y=0|\boldsymbol{W}).
\end{aligned}$$

Now, under the condition $f(Y=0|H,\boldsymbol{W})\approx f(Y=0|\boldsymbol{W})$, we have

$$\begin{aligned}
E(H|\boldsymbol{W},Y=0) &= \int H\frac{f(Y=0|H,\boldsymbol{W})f(H|\boldsymbol{W})}{\int f(Y=0|H,\boldsymbol{W})f(H|\boldsymbol{W})dh}dh \\
&\approx \int H\frac{f(Y=0|\boldsymbol{W})f(H|\boldsymbol{W})}{\int f(Y=0|\boldsymbol{W})f(H|\boldsymbol{W})dh}dh \\
&= \int H\frac{f(H|\boldsymbol{W})}{\int f(H|\boldsymbol{W})dh}dh \\
&= E(H|\boldsymbol{W}),
\end{aligned}$$

which imples $\alpha_{0j,co} \approx \alpha_{0j}$ and $\boldsymbol{\tau}_{co} \approx \boldsymbol{\tau}$. Furthermore, a small exposure effect also indicates $f(Y=0|X, \boldsymbol{Z}^*) \approx f(Y=0|\boldsymbol{Z}^*)$. Therefore,

$$\begin{aligned} f(X|\boldsymbol{H}, \boldsymbol{W}, Y=0) &= \frac{f(Y=0|X, \boldsymbol{Z}^*)f(X|\boldsymbol{H}, \boldsymbol{W})}{\int f(Y=0|X, \boldsymbol{Z}^*)f(X|\boldsymbol{H}, \boldsymbol{W})dx} \\ &\approx \frac{f(Y=0|\boldsymbol{Z}^*)f(X|\boldsymbol{H}, \boldsymbol{W})}{\int f(Y=0|\boldsymbol{Z}^*)f(X|\boldsymbol{H}, \boldsymbol{W})dx} \\ &= \frac{f(X|\boldsymbol{H}, \boldsymbol{W})}{\int f(X|\boldsymbol{H}, \boldsymbol{W})dx} \\ &= f(X|\boldsymbol{H}, \boldsymbol{W}). \end{aligned}$$

So, $f(X|\boldsymbol{H}, \boldsymbol{W}, Y=0) \approx f(X|\boldsymbol{H}, \boldsymbol{W})$.

Condition II: Rare disease prevalence. Rare disease prevalence implies $f(Y=0|H, \boldsymbol{W}) \approx 1$ and $f(Y=0|X, \boldsymbol{Z}^*) \approx 1$. Thus,

$$\begin{aligned} E(H|\boldsymbol{W}, Y=0) &= \int H \frac{f(Y=0|H, \boldsymbol{W})f(H|\boldsymbol{W})}{\int f(Y=0|H, \boldsymbol{W})f(H|\boldsymbol{W})dh} dh \\ &\approx \int H \frac{1 f(H|\boldsymbol{W})}{\int 1 f(H|\boldsymbol{W})dh} dh \\ &= \int H \frac{f(H|\boldsymbol{W})}{\int f(H|\boldsymbol{W})dh} dh \\ &= E(H|\boldsymbol{W}), \end{aligned}$$

which imples $\alpha_{0j,co} \approx \alpha_{0j}$ and $\boldsymbol{\tau}_{co} \approx \boldsymbol{\tau}$. Furthermore,

$$\begin{aligned} f(X|\boldsymbol{H}, \boldsymbol{W}, Y=0) &= \frac{f(Y=0|X, \boldsymbol{Z}^*)f(X|\boldsymbol{H}, \boldsymbol{W})}{\int f(Y=0|X, \boldsymbol{Z}^*)f(X|\boldsymbol{H}, \boldsymbol{W})dx} \\ &\approx \frac{1 f(X|\boldsymbol{H}, \boldsymbol{W})}{\int 1 f(X|\boldsymbol{H}, \boldsymbol{W})dx} \\ &= \frac{f(X|\boldsymbol{H}, \boldsymbol{W})}{\int f(X|\boldsymbol{H}, \boldsymbol{W})dx} \\ &= f(X|\boldsymbol{H}, \boldsymbol{W}). \end{aligned}$$

Therefore, $f(X|\boldsymbol{H}, \boldsymbol{W}, Y=0) \approx f(X|\boldsymbol{H}, \boldsymbol{W})$.

It follows that $E(H|\boldsymbol{W}, Y=0) \approx E(H|\boldsymbol{W})$ and $f(X|\boldsymbol{H}, \boldsymbol{W}, Y=0) \approx f(X|\boldsymbol{H}, \boldsymbol{W})$ under small exposure effects and/or rare disease prevalences.

Appendix E: Simulation Studies for Negative Variance. In this section, we implement additional simulation studies to investigate the impact on $\boldsymbol{\beta}_x$ when using the restricted method in presence of negative variance. Noting that negative variances may happen when calibration sample size and measurement errors of some laboratories are typically small. We set two studies contributing to the final analysis (i.e $M = 2$). Each study includes 1000 individuals, but only 30 controls were seleted into the calibration subset (under COCS design). Moreover, we considered three simulation scenarios for the measurement errors (1) Scenario I, the reference laboratory's measurement error was significantly smaller than those of the local laboratories; we set $\sigma_0^2 \sim U(0.25, 0.75)$ and $\sigma_d^2 \sim U(1.5, 2.5)$ for $d = 1, 2$; (2) Scenario II, the first study's measurement error was significantly smaller than those of the others; we set $\sigma_1^2 \sim U(0.25, 0.75)$ and $\sigma_d^2 \sim U(1.5, 2.5)$ for $d = 0, 2$; (3) Scenario III, all laboratories' measurement errors were typically small; we set $\sigma_d^2 \sim U(0.25, 0.75)$ for $d = 1, 2, 3$. Except the measurement errors, all the other values used in data generation were the same as those specified in Section 3.1. As follows are the data generation and estimation procedures. We at first generate Y, $\boldsymbol{W}$ and $\boldsymbol{H}$; next we used the unrestricted method to calculate $\hat{\boldsymbol{\sigma}}^2$. If the negative variance happened, we applied the restriction method to obtain the calibration parameter estimates; otherwise, we regenerated the samples, and we repeated these steps until negative variance happened. Finally, we applied the naive method, CCM and ECM to calculate $\hat{\boldsymbol{\beta}}_x$. The empirical probabilities of at least one negative variance appearance in our simulation study were about 23.1%, 24.8%, 28.0% for Scenarios I, II, III respectively.

The simulation results for Scenarios I, II, III are shown in the Supplementary Material Tables S7, S8 and S9 respectively. Among all the three scenarios, the naive method performed poorly in terms of the percent bias and coverage rate regardless of effect sizes and disease prevalences, implying the necessity of calibration. Between the two calibration methods, the percent bias of the effect estimates was generally minimized by the exact method. In the ECM, the standard sandwich variance could heavily widen the confidence interval; especially for Scenario II and III, some empirical coverage rates were 100%. However, the hessian matrix method still provided satisfactory confidence intervals under all three scenarios. This is because we optimize $L_{\boldsymbol{\sigma}^2}$ under the constrain $\boldsymbol{\sigma}^2 \geq \mathbf{0}$, instead of solving $\Psi_{\boldsymbol{\sigma}^2} = \mathbf{0}$ directly, in the presence of negative variance. Therefore, $\Psi_{\hat{\boldsymbol{\sigma}}^2} \neq \mathbf{0}$. However, the sandwich variance method mistakenly assumes $\Psi_{\hat{\boldsymbol{\sigma}}^2} = \mathbf{0}$ when taking into account the variance due to estimating $\boldsymbol{\sigma}^2$. By contrast, the hessian matrix method in the ECM considers $\boldsymbol{\sigma}^2$ as fixed values when estimating $\mathrm{var}(\hat{\boldsymbol{\beta}})$; thus the variation of $\boldsymbol{\sigma}^2$ plays a relative small effect on the final confidence interval.

In summary, the exact calibration method performed well in the presence of negative variance. We recommend using hessian matrix method to calculate the confidence interval of $\boldsymbol{\beta}_x$.

Appendix F: Proof in Section 5. Note that $H_{jk,d} = X_{jk} + \epsilon_{jk,d}$. It follows that $T(H_{jk,d}) = T(X_{jk}+\epsilon_{jk,d})$. Applying a Taylor series expansion with respect to X_{jk} and noticing $\epsilon_{jk,d}$ is small, we have $T(H_{jk,d}) \approx T(X_{jk}) + T^{'}(X_{jk})\epsilon_{jk,d}$. Thus $X_{jk} = T^{-1}(\mu_{X_{jk}|\boldsymbol{W}_{jk}} +$

$\epsilon_{x_{jk}}$), and

$$T(H_{jk,d}) \approx T(X_{jk}) + T^{'}(X_{jk})\epsilon_{jk,d}, \\ = \mu_{X_{jk}|\boldsymbol{W}_{jk}} + \epsilon_{x_{jk}} + S(\mu_{X_{jk}|\boldsymbol{W}_{jk}} + \epsilon_{x_{jk}})\epsilon_{jk,d}, \tag{A.4}$$

where $S(.) = T^{'}T^{-1}(.)$. We use a Taylor series to expand $S(\mu_{X_{jk}|\boldsymbol{W}_{jk}} + \epsilon_{x_{jk}})$ with respect to $\mu_{X_{jk}|\boldsymbol{W}_{jk}}$, which results in

$$S(\mu_{X_{jk}|\boldsymbol{W}_{jk}} + \epsilon_{x_{jk}}) = S(\mu_{X_{jk}|\boldsymbol{W}_{jk}}) + S^{'}(\mu_{X_{jk}|\boldsymbol{W}_{jk}})\epsilon_{x_{jk}} + O(\epsilon^2_{x_{jk}}), \tag{A.5}$$

where $O(\epsilon^2_{x_{jk}})$ denotes the remainder. Now substituting (A.5) to (A.6), we have

$$T(H_{jk,d}) \approx \mu_{X_{jk}|\boldsymbol{W}_{jk}} + \epsilon_{x_{jk}} + S(\mu_{X_{jk}|\boldsymbol{W}_{jk}})\epsilon_{jk,d} + S^{'}(\mu_{X_{jk}|\boldsymbol{W}_{jk}})\epsilon_{x_{jk}}\epsilon_{jk,d} + O(\epsilon^2_{x_{jk}}\epsilon_{jk,d}). \tag{A.6}$$

Again, noting $\epsilon_{jk,d}$ is small, we can omit $S^{'}(\mu_{X_{jk}|\boldsymbol{W}_{jk}})\epsilon_{x_{jk}}\epsilon_{jk,d}$ and $O(\epsilon^2_{x_{jk}}\epsilon_{jk,d})$, which lead to Equation (5.2).

Appendix G: Derivation of the Odds Ratio for the $\boldsymbol{H_{j,d}-Y}$ relationship. The association between disease outcome Y and laboratory measurement $H_{j,d}$ can be shown as

$$\begin{aligned} P(Y=1|H_{j,d},\boldsymbol{Z},\boldsymbol{W}) &= \int P(Y|X,H_{j,d},\boldsymbol{Z})f(X|H_{j,d},\boldsymbol{W})dX \\ &= \int P(Y|X,\boldsymbol{Z})f(X|H_{j,d},\boldsymbol{W})dX \\ &= \int \frac{\exp\{\beta_{0j} + \sum_{l=2}^{G} \mathbb{I}(g_{l-1} \le X < g_l))\beta_{x,l} + \boldsymbol{\beta}_z^T\boldsymbol{Z}\}}{1+\exp(\beta_{0j} + \sum_{l=2}^{G} \mathbb{I}(g_{l-1} \le X < g_l))\beta_{x,l} + \boldsymbol{\beta}_z^T\boldsymbol{Z})} f(X|H_{j,d},\boldsymbol{W})dX \\ &= \sum_{l=2}^{G} \frac{\exp\{\beta_{0j} + \beta_{x,l} + \boldsymbol{\beta}_z^T\boldsymbol{Z}\}}{1+\exp(\beta_{0j} + \beta_{x,l} + \boldsymbol{\beta}_z^T\boldsymbol{Z})} P(g_{l-1} \le X < g_l|H_{j,d},\boldsymbol{W}), \end{aligned}$$

where the notations were defined in section 2.1 (some subscripts are suppressed for brevity). It follows that the Odds Ratio for one unit increase in $H_{j,d}$ is

$$\begin{aligned} \mathrm{OR}_{H_{j,d}}(h) &= \frac{\mathrm{Odds}(H_{j,d}=h+1|\boldsymbol{Z},\boldsymbol{W})}{\mathrm{Odds}(H_{j,d}=h|\boldsymbol{Z},\boldsymbol{W})} \\ &= \frac{P(Y=1|H_{j,d}=h+1,\boldsymbol{Z},\boldsymbol{W})/(1-P(Y=1|H_{j,d}=h+1,\boldsymbol{Z},\boldsymbol{W}))}{P(Y=1|H_{j,d}=h,\boldsymbol{Z},\boldsymbol{W})/(1-P(Y=1|H_{j,d}=h,\boldsymbol{Z},\boldsymbol{W}))} \\ &= \frac{[\sum_{l=2}^{G}\Delta_{j,k}(h+1,l)]/[1-\sum_{l=2}^{G}\Delta_{j,k}(h+1,l)]}{[\sum_{l=2}^{G}\Delta_{j,k}(h,l)]/[1-\sum_{l=2}^{G}\Delta_{j,k}(h,l)]}, \end{aligned}$$

where $\Delta_{j,k}(h,l) = \frac{\exp\{\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z}\}}{1+\exp(\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z})} P(g_{l-1} \le X < g_l|H_{j,d}=h,\boldsymbol{W})$. Note that $P(g_{l-1} \le X < g_l|H_{j,d}=h,\boldsymbol{W})$ can be estimated based on (2.11).

Below we show that, if the disease prevalence is low, the odds for $H_{j,d}$ conditional on $\boldsymbol{Z}$ and $\boldsymbol{W}$ can be approximated by

$$\text{Odds}(H_{j,d}=h|\boldsymbol{Z},\boldsymbol{W}) := \frac{P(Y=1|H_{j,d}=h,\boldsymbol{Z},\boldsymbol{W})}{1-P(Y=1|H_{j,d},\boldsymbol{Z},\boldsymbol{W})} \approx \sum_{l=2}^{G} \exp(\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z})P(g_{l-1}\le X<g_l|H_{j,d}=h,\boldsymbol{W}). \tag{A.7}$$

Under the rare disease prevalence assumption, $P(Y=1|H_{j,d},\boldsymbol{Z},\boldsymbol{W})\approx 0$ and $\text{Odds}(H_{j,d}=h|\boldsymbol{Z},\boldsymbol{W}) \approx P(Y=1|H_{j,d}=h,\boldsymbol{Z},\boldsymbol{W}) = \int P(Y=1|X,\boldsymbol{Z})f(X|H_{j,d}=h,\boldsymbol{W})dX$, and $P(Y=1|X,\boldsymbol{Z}) \approx \exp(\beta_{0j}+\sum_{l=2}^{G}\mathbb{I}(g_{l-1}\le X<g_l)\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z})$. It follows that $\text{Odds}(H_{j,d}=h|\boldsymbol{Z},\boldsymbol{W}) \approx \sum_{l=2}^{G}\exp(\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z})P(g_{l-1}\le X<g_l|H_{j,d}=h,\boldsymbol{W})$. Given approximation (A.7), the odds ratio for one unit increase in $H_{j,d}$ can be approximated by

$$\begin{aligned}
\text{OR}_{H_{j,d}}(h) &= \frac{\text{Odds}(H_{j,d}=h+1|\boldsymbol{Z},\boldsymbol{W})}{\text{Odds}(H_{j,d}=h|\boldsymbol{Z},\boldsymbol{W})} \\
&= \frac{\sum_{l=2}^{G}\exp(\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z})P(g_{l-1}\le X<g_l|H_{j,d}=h+1,\boldsymbol{W})}{\sum_{l=2}^{G}\exp(\beta_{0j}+\beta_{x,l}+\boldsymbol{\beta}_z^T\boldsymbol{Z})P(g_{l-1}\le X<g_l|H_{j,d}=h,\boldsymbol{W})} \\
&= \frac{\sum_{l=2}^{G}\exp(\beta_{x,l})P(g_{l-1}\le X<g_l|H_{j,d}=h+1,\boldsymbol{W})}{\sum_{l=2}^{G}\exp(\beta_{x,l})P(g_{l-1}\le X<g_l|H_{j,d}=h,\boldsymbol{W})}.
\end{aligned} \tag{A.8}$$

Substituting the estimated parameters in section 2 into $\text{OR}_{H_{j,d}}(h)$ leads to $\widehat{\text{OR}}_{H_{j,d}}(h)$.

The variance of $\widehat{\text{OR}}_{H_{j,d}}(h)$ can be estimated using the delta method, that is, $\text{var}(\widehat{\text{OR}}_{H_{j,d}}(h)) \approx [\frac{\partial \text{OR}_{H_{j,d}}(h)}{\partial \boldsymbol{\kappa}}|_{\boldsymbol{\kappa}=\hat{\boldsymbol{\kappa}}}]^T \text{var}(\hat{\boldsymbol{\kappa}}) \frac{\partial \text{OR}_{H_{j,d}}(h)}{\partial \boldsymbol{\kappa}}|_{\boldsymbol{\kappa}=\hat{\boldsymbol{\kappa}}}$, where $\boldsymbol{\kappa}=[\boldsymbol{\beta},\sigma_x^2,\sigma_d^2]^T$, and estimation of $\text{var}(\hat{\boldsymbol{\kappa}})$ is discussed in section 2.5 and 2.6.

SUPPLEMENTARY TABLES

TABLE S1. *Comparison of operating characteristics over different percentages of subjects involved in the calibration sub-study for naive ($\hat{\beta}^{(N)}$), cut-off calibration ($\hat{\beta}^{(C)}$) and exact calibration ($\hat{\beta}^{(E)}$) methods under the COCS design, disease prevalence 25%, and $\sigma_d^2 \sim Unif(1.5, 2.5)$ for $d = 0, \ldots, 5$.*

Percentage calib. study	$\beta_{x,2}$	Percent Bias $\hat{\beta}^{(N)}_{x,2}$	Percent Bias $\hat{\beta}^{(C)}_{x,2}$	Percent Bias $\hat{\beta}^{(E)}_{x,2}$	MSE $\hat{\beta}^{(N)}_{x,2}$	MSE $\hat{\beta}^{(C)}_{x,2}$	MSE $\hat{\beta}^{(E)}_{x,2}$	SE $\hat{\beta}^{(N)}_{x,2}$	SE $\hat{\beta}^{(C)}_{x,2}$	SE $\hat{\beta}^{(E)}_{x,2}$	Coverage Rate $\hat{\beta}^{(N)}_{x,2}$	Coverage Rate $\hat{\beta}^{(C)}_{x,2}$	Coverage Rate $\hat{\beta}^{(E)}_{x,2}$	Coverage Rate $\hat{\beta}^{(E*)}_{x,2}$
3%	$\frac{1}{2}\log(1.5)$	-24.9	-15.5	0.0	0.009	0.007	0.018	0.081	0.080	0.133	90.8	94.2	97.2	97.0
	$\frac{1}{2}\log(2)$	-23.3	-14.6	-0.3	0.014	0.009	0.019	0.085	0.082	0.139	84.1	90.7	96.2	95.6
	$\frac{1}{2}\log(3)$	-23.5	-14.7	0.1	0.024	0.014	0.022	0.088	0.088	0.147	66.6	81.6	95.5	94.8
10%	$\frac{1}{2}\log(1.5)$	-30.3	-13.2	-2.6	0.011	0.007	0.019	0.083	0.080	0.138	88.0	93.4	95.0	94.9
	$\frac{1}{2}\log(2)$	-26.1	-13.8	-0.4	0.015	0.009	0.019	0.085	0.084	0.138	80.6	89.8	96.2	95.8
	$\frac{1}{2}\log(3)$	-24.6	-13.1	-0.2	0.026	0.012	0.022	0.086	0.085	0.149	66.1	86.3	94.9	94.4
30%	$\frac{1}{2}\log(1.5)$	-40.8	-6.5	-2.1	0.014	0.007	0.018	0.084	0.082	0.135	81.9	94.4	95.8	95.7
	$\frac{1}{2}\log(2)$	-31.8	-8.6	0.8	0.019	0.008	0.020	0.082	0.084	0.141	74.7	92.9	96.6	96.0
	$\frac{1}{2}\log(3)$	-28.3	-11.6	-1.2	0.032	0.011	0.020	0.087	0.082	0.140	55.6	89.7	95.7	95.6

Percentage calib. study	$\beta_{x,3}$	Percent Bias $\hat{\beta}^{(N)}_{x,3}$	Percent Bias $\hat{\beta}^{(C)}_{x,3}$	Percent Bias $\hat{\beta}^{(E)}_{x,3}$	MSE $\hat{\beta}^{(N)}_{x,3}$	MSE $\hat{\beta}^{(C)}_{x,3}$	MSE $\hat{\beta}^{(E)}_{x,3}$	SE $\hat{\beta}^{(N)}_{x,3}$	SE $\hat{\beta}^{(C)}_{x,3}$	SE $\hat{\beta}^{(E)}_{x,3}$	Coverage Rate $\hat{\beta}^{(N)}_{x,3}$	Coverage Rate $\hat{\beta}^{(C)}_{x,3}$	Coverage Rate $\hat{\beta}^{(E)}_{x,3}$	Coverage Rate $\hat{\beta}^{(E*)}_{x,3}$
3%	$\log(1.5)$	-24.5	-18.4	-1.4	0.016	0.013	0.009	0.079	0.084	0.097	75.3	84.8	95.2	94.8
	$\log(2)$	-24.5	-17.8	-0.8	0.036	0.022	0.010	0.081	0.085	0.100	44.2	68.4	94.2	94.0
	$\log(3)$	-24.8	-18.0	-0.4	0.081	0.046	0.011	0.080	0.085	0.105	8.3	36.1	95.9	93.7
10%	$\log(1.5)$	-24.1	-16.9	0.1	0.016	0.012	0.009	0.079	0.084	0.097	77.8	86.2	93.7	93.6
	$\log(2)$	-24.7	-17.9	-0.6	0.037	0.023	0.011	0.086	0.090	0.105	42.4	67.2	93.4	93.1
	$\log(3)$	-24.7	-17.5	0.1	0.080	0.044	0.011	0.080	0.086	0.104	8.1	37.1	94.7	94.0
30%	$\log(1.5)$	-24.1	-17.6	-0.9	0.016	0.012	0.009	0.080	0.085	0.097	76.1	83.5	93.9	93.8
	$\log(2)$	-23.6	-16.9	0.3	0.033	0.021	0.009	0.079	0.084	0.097	46.6	70.4	94.5	94.4
	$\log(3)$	-24.5	-17.7	-0.4	0.079	0.045	0.010	0.079	0.082	0.098	8.1	34.2	95.2	95.1

NOTE: Percent bias and MSE were computed by averaging $(\hat{\beta} - \beta)/\beta$ and $(\beta - \hat{\beta})^2$ over 1000 simulations. Standard error (SE) is the square root of the empirical variance over all replicates. Coverage rate represents the coverage of a 95% confidence interval. We used the sandwich variances for the confidence intervals of the naive and cut-off calibration methods. For the exact calibration method, we applied both the sandwich and hessian matrix approaches for estimating the variances in the coverage rate evaluation, denoting by $\hat{\beta}^{(E)}$ and $\hat{\beta}^{(E*)}$ in the "Coverage Rate" columns respectively.

Table S2. *Comparison of operating characteristics over different magnitude of σ_x^2 for naive ($\hat{\beta}^{(N)}$), cut-off calibration ($\hat{\beta}^{(C)}$) and exact calibration ($\hat{\beta}^{(E)}$) methods under the COCS design; the ICC corresponding to $\sigma_x^2 = 6, 3, 0.5$ ranged $[70.6\%, 80.0\%]$, $[54.5\%, 66.7\%]$, and $[16.7\%, 25.0\%]$ respectively.*

σ_x^2	$\beta_{x,2}$	Percent Bias			MSE			SE			Coverage Rate			
		$\hat{\beta}_{x,2}^{(N)}$	$\hat{\beta}_{x,2}^{(C)}$	$\hat{\beta}_{x,2}^{(E)}$	$\hat{\beta}_{x,2}^{(N)}$	$\hat{\beta}_{x,2}^{(C)}$	$\hat{\beta}_{x,2}^{(E)}$	$\hat{\beta}_{x,2}^{(N)}$	$\hat{\beta}_{x,2}^{(C)}$	$\hat{\beta}_{x,2}^{(E)}$	$\hat{\beta}_{x,2}^{(N)}$	$\hat{\beta}_{x,2}^{(C)}$	$\hat{\beta}_{x,2}^{(E)}$	$\hat{\beta}_{x,2}^{(E*)}$
6	$\frac{1}{2}\log(1.5)$	-28.7	-12.8	0.1	0.010	0.007	0.023	0.083	0.081	0.150	89.7	93.6	96.3	95.2
	$\frac{1}{2}\log(2)$	-26.8	-14.3	-0.5	0.015	0.010	0.024	0.083	0.085	0.156	80.5	90.8	95.9	95.5
	$\frac{1}{2}\log(3)$	-24.7	-14.3	0.1	0.026	0.013	0.026	0.087	0.084	0.160	63.1	85.7	96.7	95.4
3	$\frac{1}{2}\log(1.5)$	-30.3	-13.2	-2.6	0.011	0.007	0.019	0.083	0.080	0.138	88.0	93.4	95.0	94.9
	$\frac{1}{2}\log(2)$	-26.1	-13.8	-0.4	0.015	0.009	0.019	0.085	0.084	0.138	80.6	89.8	96.2	95.8
	$\frac{1}{2}\log(3)$	-24.6	-13.1	-0.2	0.026	0.012	0.022	0.086	0.085	0.149	66.1	86.3	94.9	94.4
0.5	$\frac{1}{2}\log(1.5)$	-29.4	-8.4	0.9	0.010	0.007	0.011	0.082	0.081	0.105	89.1	94.5	95.1	95.1
	$\frac{1}{2}\log(2)$	-24.8	-8.4	0.5	0.014	0.008	0.012	0.083	0.087	0.109	83.5	91.6	94.1	94.0
	$\frac{1}{2}\log(3)$	-24.2	-8.3	-0.0	0.026	0.009	0.012	0.090	0.084	0.110	64.3	91.9	95.6	95.4

σ_x^2	$\beta_{x,3}$	Percent Bias			MSE			SE			Coverage Rate			
		$\hat{\beta}_{x,3}^{(N)}$	$\hat{\beta}_{x,3}^{(C)}$	$\hat{\beta}_{x,3}^{(E)}$	$\hat{\beta}_{x,3}^{(N)}$	$\hat{\beta}_{x,3}^{(C)}$	$\hat{\beta}_{x,3}^{(E)}$	$\hat{\beta}_{x,3}^{(N)}$	$\hat{\beta}_{x,3}^{(C)}$	$\hat{\beta}_{x,3}^{(E)}$	$\hat{\beta}_{x,3}^{(N)}$	$\hat{\beta}_{x,3}^{(C)}$	$\hat{\beta}_{x,3}^{(E)}$	$\hat{\beta}_{x,3}^{(E*)}$
6	$\log(1.5)$	-24.6	-19.9	-0.7	0.016	0.013	0.010	0.078	0.083	0.100	76.0	83.9	94.8	94.5
	$\log(2)$	-24.2	-19.3	0.2	0.035	0.026	0.012	0.082	0.088	0.108	43.2	61.8	93.5	92.6
	$\log(3)$	-24.9	-19.8	0.3	0.082	0.055	0.014	0.083	0.089	0.119	8.8	28.5	94.4	92.2
3	$\log(1.5)$	-24.1	-16.9	0.1	0.016	0.012	0.009	0.079	0.084	0.097	77.8	86.2	93.7	93.6
	$\log(2)$	-24.7	-17.9	-0.6	0.037	0.023	0.011	0.086	0.090	0.105	42.4	67.2	93.4	93.1
	$\log(3)$	-24.7	-17.5	0.1	0.080	0.044	0.011	0.080	0.086	0.104	8.1	37.1	94.7	94.0
0.5	$\log(1.5)$	-24.6	-10.6	-0.5	0.016	0.008	0.008	0.078	0.082	0.087	75.2	91.8	95.4	95.4
	$\log(2)$	-24.5	-9.9	0.2	0.035	0.012	0.008	0.077	0.084	0.091	42.6	86.0	94.5	94.5
	$\log(3)$	-24.5	-10.1	0.2	0.078	0.019	0.008	0.078	0.081	0.089	8.4	72.8	95.7	95.6

NOTE: Percent bias and MSE were computed by averaging $(\hat{\beta} - \beta)/\beta$ and $(\beta - \hat{\beta})^2$ over 1000 simulations. Standard error (SE) is the square root of the empirical variance over all replicates. Coverage rate represents the coverage of a 95% confidence interval. We used the sandwich variances for the confidence intervals of the naive and cut-off calibration methods. For the exact calibration method, we applied both the sandwich and hessian matrix approaches for estimating the variances in the coverage rate evaluation, denoting by $\hat{\beta}^{(E)}$ and $\hat{\beta}^{(E*)}$ in the "Coverage Rate" columns respectively.

TABLE S3. *Comparison of operating characteristics over different magnitude of measurement errors (i.e σ_d^2 for $d = 0, \ldots, 5$) for naive ($\hat{\beta}^{(N)}$), cut-off calibration ($\hat{\beta}^{(C)}$) and exact calibration ($\hat{\beta}^{(E)}$) methods under the COCS design; the ICC range corresponding to $\sigma_d^2 \sim$ Unif$(3, 4)$, Unif$(1.5, 2.5)$ and Unif$(0.25, 0.75)$ ranged* [42.9%, 50.0%], [54.5%, 66.7%] *and* [80.0%, 92.3%] *respectively.*

σ_d^2		Percent Bias			MSE			SE			Coverage Rate			
	$\beta_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(E*)}_{x,2}$
U(3,4)	$\frac{1}{2}\log(1.5)$	-39.6	-14.4	-2.0	0.013	0.008	0.025	0.084	0.082	0.159	84.6	93.6	95.2	95.0
	$\frac{1}{2}\log(2)$	-35.7	-15.3	1.1	0.023	0.010	0.028	0.088	0.086	0.167	68.7	89.0	94.0	93.8
	$\frac{1}{2}\log(3)$	-32.9	-15.5	1.1	0.040	0.014	0.026	0.088	0.084	0.161	43.6	82.8	96.2	95.9
U(1.5,2.5)	$\frac{1}{2}\log(1.5)$	-30.3	-13.2	-2.6	0.011	0.007	0.019	0.083	0.080	0.138	88.0	93.4	95.0	94.9
	$\frac{1}{2}\log(2)$	-26.1	-13.8	-0.4	0.015	0.009	0.019	0.085	0.084	0.138	80.6	89.8	96.2	95.8
	$\frac{1}{2}\log(3)$	-24.6	-13.1	-0.2	0.026	0.012	0.022	0.086	0.085	0.149	66.1	86.3	94.9	94.4
U(0.25,0.75)	$\frac{1}{2}\log(1.5)$	-13.8	-9.7	-1.2	0.008	0.007	0.011	0.083	0.082	0.105	93.2	94.0	95.9	94.9
	$\frac{1}{2}\log(2)$	-11.2	-8.0	0.1	0.008	0.008	0.011	0.082	0.083	0.106	93.1	94.2	95.8	95.1
	$\frac{1}{2}\log(4)$	-11.0	-8.1	0.0	0.011	0.010	0.013	0.087	0.087	0.113	89.4	92.4	96.5	95.0

σ_d^2		Percent Bias			MSE			SE			Coverage Rate			
	$\beta_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(E*)}_{x,3}$
U(3,4)	$\log(1.5)$	-31.4	-19.3	0.8	0.022	0.013	0.010	0.077	0.083	0.098	62.5	86.4	95.0	94.9
	$\log(2)$	-32.4	-20.9	-0.2	0.057	0.028	0.011	0.079	0.084	0.104	17.9	59.5	93.8	93.8
	$\log(3)$	-32.4	-20.4	0.7	0.133	0.058	0.012	0.082	0.088	0.108	0.6	25.9	95.0	94.3
U(1.5,2.5)	$\log(1.5)$	-24.1	-16.9	0.1	0.016	0.012	0.009	0.079	0.084	0.097	77.8	86.2	93.7	93.6
	$\log(2)$	-24.7	-17.9	-0.6	0.037	0.023	0.011	0.086	0.090	0.105	42.4	67.2	93.4	93.1
	$\log(3)$	-24.7	-17.5	0.1	0.080	0.044	0.011	0.080	0.086	0.104	8.1	37.1	94.7	94.0
U(0.25,0.75)	$\log(1.5)$	-12.0	-10.7	-0.8	0.009	0.008	0.008	0.081	0.081	0.090	91.6	91.5	95.1	94.8
	$\log(2)$	-11.8	-10.2	-0.1	0.013	0.012	0.008	0.081	0.083	0.090	80.5	84.6	96.7	96.0
	$\log(3)$	-12.2	-10.6	-0.4	0.025	0.021	0.010	0.083	0.085	0.098	62.3	70.6	95.2	93.7

NOTE: Percent bias and MSE were computed by averaging $(\hat{\beta} - \beta)/\beta$ and $(\beta - \hat{\beta})^2$ over 1000 simulations. Standard error (SE) is the square root of the empirical variance over all replicates. Coverage rate represents the coverage of a 95% confidence interval. We used the sandwich variances for the confidence intervals of the naive and cut-off calibration methods. For the exact calibration method, we applied both the sandwich and hessian matrix approaches for estimating the variances in the coverage rate evaluation, denoting by $\hat{\beta}^{(E)}$ and $\hat{\beta}^{(E*)}$ in the "Coverage Rate" columns respectively.

TABLE S4. *Comparison of operating characteristics under the COCS design for naive ($\hat{\beta}^{(N)}$), cut-off calibration ($\hat{\beta}^{(C)}$) and exact calibration ($\hat{\beta}^{(E)}$) methods; assuming $\boldsymbol{Z} = W$ is included in model (3.1) with regression coefficient $\beta_z = 1$.*

Disease		Percent Bias			MSE			SE			Coverage Rate			
Prevalence	$\beta_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(E*)}_{x,2}$
5%	$\frac{1}{2}\log(1.5)$	-55.8	-5.5	-4.5	0.050	0.059	0.218	0.193	0.243	0.467	91.9	95.2	95.5	95.4
	$\frac{1}{2}\log(2)$	-49.9	-8.7	4.2	0.069	0.066	0.221	0.198	0.256	0.470	85.9	93.7	96.4	95.7
	$\frac{1}{2}\log(3)$	-45.9	-17.7	2.0	0.105	0.070	0.222	0.204	0.246	0.472	75.5	93.1	96.8	96.4
25%	$\frac{1}{2}\log(1.5)$	-55.8	-1.7	10.3	0.047	0.042	0.165	0.184	0.206	0.406	90.0	95.4	95.7	94.8
	$\frac{1}{2}\log(2)$	-48.8	-13.7	3.7	0.062	0.048	0.165	0.181	0.215	0.406	84.6	93.3	96.4	96.1
	$\frac{1}{2}\log(3)$	-45.0	-19.2	1.7	0.097	0.058	0.181	0.190	0.218	0.425	71.9	90.4	96.1	95.1
50%	$\frac{1}{2}\log(1.5)$	-55.9	-5.4	-0.5	0.043	0.035	0.139	0.175	0.186	0.373	90.7	95.9	96.6	96.1
	$\frac{1}{2}\log(2)$	-50.4	-16.4	-2.7	0.063	0.042	0.157	0.181	0.198	0.396	82.3	93.0	94.6	94.4
	$\frac{1}{2}\log(3)$	-46.5	-21.7	0.1	0.096	0.052	0.143	0.176	0.195	0.379	69.4	89.6	96.3	94.9

Disease		Percent Bias			MSE			SE			Coverage Rate			
Prevalence	$\beta_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(E*)}_{x,3}$
5%	$\log(1.5)$	-41.9	-21.1	1.8	0.065	0.074	0.112	0.191	0.258	0.335	86.7	94.3	96.0	95.0
	$\log(2)$	-41.8	-21.8	3.0	0.122	0.091	0.114	0.195	0.262	0.337	68.4	91.9	96.6	95.8
	$\log(3)$	-42.4	-26.8	0.1	0.256	0.156	0.126	0.200	0.263	0.355	33.2	79.5	96.4	95.5
25%	$\log(1.5)$	-41.4	-25.0	1.7	0.068	0.069	0.099	0.200	0.243	0.314	84.1	92.6	94.0	93.8
	$\log(2)$	-41.2	-27.8	-0.6	0.117	0.097	0.094	0.188	0.245	0.306	67.1	85.3	95.4	95.0
	$\log(3)$	-39.5	-26.9	2.0	0.227	0.149	0.102	0.196	0.249	0.319	36.7	75.8	95.4	93.8
50%	$\log(1.5)$	-40.0	-26.6	2.2	0.062	0.064	0.085	0.189	0.230	0.291	86.9	93.3	95.7	95.4
	$\log(2)$	-38.6	-26.6	2.6	0.107	0.088	0.089	0.189	0.233	0.298	68.9	86.9	95.4	95.4
	$\log(3)$	-39.0	-27.9	1.8	0.218	0.149	0.095	0.188	0.235	0.308	36.9	72.6	95.9	95.3

NOTE: Percent bias and MSE were computed by averaging $(\hat{\beta} - \beta)/\beta$ and $(\beta - \hat{\beta})^2$ over 1000 simulations. Standard error (SE) is the square root of the empirical variance over all replicates. Coverage rate represents the coverage of a 95% confidence interval. We used the sandwich variances for the confidence intervals of the naive and cut-off calibration methods. For the exact calibration method, we applied both the sandwich and hessian matrix approaches for estimating the variances in the coverage rate evaluation, denoting by $\hat{\beta}^{(E)}$ and $\hat{\beta}^{(E*)}$ in the "Coverage Rate" columns respectively.

TABLE S5. *Comparison of operating characteristics over $\alpha_{0j} = 2j$ (where $j = 1, \ldots, 5$) for naive ($\hat{\beta}^{(N)}$), cut-off calibration ($\hat{\beta}^{(C)}$) and exact calibration ($\hat{\beta}^{(E)}$) methods under the COCS design.*

Disease		Percent Bias			MSE			SE			Coverage Rate			
Prevalence	$\beta_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(E*)}_{x,2}$
5%	$\frac{1}{2}\log(1.5)$	-34.8	-17.0	-13.2	0.041	0.040	0.103	0.189	0.198	0.319	94.4	94.9	96.4	96.3
	$\frac{1}{2}\log(2)$	-30.3	-16.4	-9.1	0.050	0.043	0.106	0.197	0.199	0.324	91.4	93.6	95.0	94.4
	$\frac{1}{2}\log(3)$	-27.2	-13.2	-5.9	0.059	0.046	0.100	0.192	0.202	0.314	87.0	93.5	96.2	96.1
25%	$\frac{1}{2}\log(1.5)$	-30.9	-14.2	1.4	0.013	0.010	0.022	0.097	0.097	0.150	89.4	94.5	95.4	95.1
	$\frac{1}{2}\log(2)$	-29.0	-15.3	-1.8	0.020	0.013	0.024	0.099	0.100	0.155	82.3	91.1	94.7	94.4
	$\frac{1}{2}\log(3)$	-27.7	-14.2	0.6	0.033	0.017	0.025	0.099	0.103	0.158	64.7	87.2	94.1	93.7
50%	$\frac{1}{2}\log(1.5)$	-32.0	-16.9	-1.2	0.011	0.008	0.015	0.081	0.081	0.123	87.6	94.4	96.3	96.2
	$\frac{1}{2}\log(2)$	-30.1	-17.1	-1.2	0.018	0.011	0.017	0.082	0.086	0.131	76.2	90.6	95.4	95.3
	$\frac{1}{2}\log(3)$	-30.6	-16.4	0.4	0.035	0.015	0.017	0.080	0.086	0.131	48.4	82.3	94.9	94.9
Disease		Percent Bias			MSE			SE			Coverage Rate			
Prevalence	$\beta_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(E*)}_{x,3}$
5%	$\log(1.5)$	-27.2	-15.0	2.3	0.048	0.050	0.058	0.190	0.215	0.240	92.7	94.1	95.4	95.4
	$\log(2)$	-27.8	-17.6	0.3	0.076	0.062	0.062	0.197	0.217	0.250	83.2	91.0	93.8	94.3
	$\log(3)$	-29.4	-18.0	-0.9	0.140	0.085	0.060	0.190	0.214	0.244	58.2	84.8	94.5	94.3
25%	$\log(1.5)$	-29.1	-18.1	-0.2	0.023	0.017	0.015	0.097	0.108	0.123	76.5	89.3	95.2	95.0
	$\log(2)$	-28.6	-17.1	0.3	0.049	0.026	0.016	0.098	0.111	0.127	47.0	80.3	94.5	94.5
	$\log(3)$	-29.0	-17.5	0.9	0.111	0.050	0.018	0.100	0.113	0.136	11.1	59.2	94.3	94.0
50%	$\log(1.5)$	-28.2	-16.6	0.6	0.020	0.013	0.012	0.083	0.092	0.108	72.4	89.5	95.6	95.6
	$\log(2)$	-28.9	-17.5	-0.4	0.047	0.024	0.012	0.086	0.095	0.111	35.2	75.8	94.7	94.5
	$\log(3)$	-29.8	-18.0	-0.5	0.115	0.048	0.013	0.084	0.095	0.112	2.8	46.4	95.7	95.2

NOTE: Percent bias and MSE were computed by averaging $(\hat{\beta} - \beta)/\beta$ and $(\beta - \hat{\beta})^2$ over 1000 simulations. Standard error (SE) is the square root of the empirical variance over all replicates. Coverage rate represents the coverage of a 95% confidence interval. We used the sandwich variances for the confidence intervals of the naive and cut-off calibration methods. For the exact calibration method, we applied both the sandwich and hessian matrix approaches for estimating the variances in the coverage rate evaluation, denoting by $\hat{\beta}^{(E)}$ and $\hat{\beta}^{(E*)}$ in the "Coverage Rate" columns respectively.

TABLE S6. *Comparison of operating characteristics under a COCS design for the naive method ($\hat{\beta}^{(N)}$), the Full Calibration method in Sloan et al. (2019) ($\hat{\beta}^{(F)}$), cut-off calibration method ($\hat{\beta}^{(C)}$) and exact calibration method ($\hat{\beta}^{(E)}$).*

Disease		Percent Bias				MSE				SE				Coverage Rate			
Prev.	$\beta_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(R)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(F)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(F)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(F)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$
5%	$\frac{1}{2}\log(1.2)$	-33.4	-15.9	-5.6	-3.1	0.028	0.027	0.026	0.079	0.164	0.163	0.162	0.280	95.2	94.6	94.8	96.3
	$\frac{1}{2}\log(1.5)$	-25.2	-16.5	-13.1	-2.5	0.033	0.031	0.029	0.085	0.173	0.172	0.169	0.292	95.1	94.4	93.9	96.3
	$\frac{1}{2}\log(2)$	-25.5	-15.5	-11.4	-1.1	0.037	0.032	0.032	0.090	0.172	0.172	0.173	0.301	92.5	94.7	94.9	95.6
	$\frac{1}{2}\log(3)$	-22.2	-13.3	-13.6	-3.8	0.049	0.037	0.038	0.104	0.186	0.179	0.179	0.321	88.1	92.6	91.9	96.4
25%	$\frac{1}{2}\log(1.2)$	-36.1	-23.4	-5.6	4.9	0.008	0.007	0.007	0.019	0.081	0.081	0.081	0.137	92.7	94.0	95.4	96.0
	$\frac{1}{2}\log(1.5)$	-30.3	-17.3	-13.2	-2.6	0.011	0.007	0.007	0.019	0.083	0.079	0.080	0.138	88.0	93.3	93.4	95.0
	$\frac{1}{2}\log(2)$	-26.1	-18.0	-13.8	-0.4	0.015	0.01	0.009	0.019	0.085	0.081	0.084	0.138	80.6	87.7	89.8	96.2
	$\frac{1}{2}\log(3)$	-24.6	-16.8	-13.1	-0.2	0.026	0.016	0.012	0.022	0.086	0.084	0.085	0.149	66.1	80.3	86.3	94.9
50%	$\frac{1}{2}\log(1.2)$	-42.3	-25.3	-4.8	2.9	0.006	0.005	0.005	0.014	0.069	0.069	0.069	0.116	92.7	92.9	95.3	95.4
	$\frac{1}{2}\log(1.5)$	-33.6	-18.5	-12.1	-1.4	0.009	0.006	0.005	0.012	0.067	0.071	0.067	0.110	85.5	90.7	93.9	97.3
	$\frac{1}{2}\log(2)$	-28.3	-18.7	-13.7	0.1	0.015	0.009	0.007	0.013	0.071	0.067	0.069	0.115	71.1	90.7	89.3	95.5
	$\frac{1}{2}\log(3)$	-26.8	-18.2	-15.5	0.5	0.027	0.015	0.012	0.014	0.070	0.068	0.067	0.118	45.9	68.9	77.3	95.6

Disease		Percent Bias				MSE				SE				Coverage Rate			
Prev.	$\beta_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(F)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(F)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(F)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(F)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$
5%	$\log(1.2)$	-23.3	-16.9	-15.7	2.1	0.026	0.027	0.027	0.035	0.157	0.162	0.163	0.187	94.3	95.9	95.1	96.2
	$\log(1.5)$	-23.8	-19.9	-17.1	0.9	0.035	0.035	0.033	0.038	0.159	0.168	0.167	0.194	90.0	91.8	93.3	95.1
	$\log(2)$	-24.3	-19.3	-17.3	0.7	0.053	0.047	0.044	0.041	0.156	0.169	0.172	0.202	81.1	85.9	87.0	94.9
	$\log(3)$	-24.7	-20.1	-18.1	-0.2	0.100	0.080	0.068	0.044	0.161	0.176	0.169	0.210	57.7	71.6	75.7	95.5
25%	$\log(1.2)$	-20.8	-21.3	-14.3	3.8	0.007	0.009	0.007	0.009	0.077	0.084	0.081	0.094	93.0	92.6	93.5	95.3
	$\log(1.5)$	-24.1	-20.4	-16.9	0.1	0.016	0.014	0.012	0.009	0.079	0.084	0.084	0.097	77.8	81.1	86.2	93.7
	$\log(2)$	-24.7	-20.2	-17.9	-0.6	0.037	0.027	0.023	0.011	0.086	0.086	0.090	0.105	42.4	60.8	67.2	93.4
	$\log(3)$	-24.7	-20.6	-17.5	0.1	0.080	0.059	0.044	0.011	0.080	0.088	0.086	0.104	8.1	24.9	37.1	94.7
50%	$\log(1.2)$	-25.7	-21.2	-19.1	-2.9	0.007	0.006	0.006	0.007	0.066	0.070	0.070	0.081	89.9	92.2	92.8	95.5
	$\log(1.5)$	-24.2	-20.8	-17.2	-0.6	0.014	0.012	0.010	0.007	0.066	0.073	0.070	0.081	70.8	77.9	83.5	95.5
	$\log(2)$	-23.7	-20.0	-16.9	0.3	0.032	0.025	0.019	0.008	0.070	0.074	0.074	0.087	33.5	51.1	62.8	93.6
	$\log(3)$	-24.3	-20.4	-17.5	-0.0	0.076	0.056	0.042	0.008	0.068	0.075	0.073	0.087	3.1	14.7	24.6	95.4

NOTE: Percent bias and MSE were computed by averaging $(\hat{\beta}-\beta)/\beta$ and $(\beta-\hat{\beta})^2$ over 1000 simulations. Standard error (SE) is the square root of the empirical variance over all replicates. Coverage rate represents the coverage of a 95% confidence interval. We used the sandwich variances for the confidence intervals of all methods.

Table S7. *Comparison of operating characteristics in presence of negative variances for naive ($\hat{\beta}^{(N)}$), cut-off calibration ($\hat{\beta}^{(C)}$) and exact calibration ($\hat{\beta}^{(E)}$) methods, under Scenario I in Appendix A, i.e $\sigma_0^2 \sim U(0.25, 0.75)$ and $\sigma_d^2 \sim U(1.5, 2.5)$ for $d = 1, 2$.*

Disease Prevalence	$\beta_{x,2}$	Percent Bias $\hat{\beta}^{(N)}_{x,2}$	Percent Bias $\hat{\beta}^{(C)}_{x,2}$	Percent Bias $\hat{\beta}^{(E)}_{x,2}$	MSE $\hat{\beta}^{(N)}_{x,2}$	MSE $\hat{\beta}^{(C)}_{x,2}$	MSE $\hat{\beta}^{(E)}_{x,2}$	SE $\hat{\beta}^{(N)}_{x,2}$	SE $\hat{\beta}^{(C)}_{x,2}$	SE $\hat{\beta}^{(E)}_{x,2}$	Coverage Rate $\hat{\beta}^{(N)}_{x,2}$	Coverage Rate $\hat{\beta}^{(C)}_{x,2}$	Coverage Rate $\hat{\beta}^{(E)}_{x,2}$	Coverage Rate $\hat{\beta}^{(E*)}_{x,2}$
5%	$\frac{1}{2}\log(1.5)$	-26.6	-17.1	-16.9	0.077	0.073	0.221	0.273	0.268	0.469	94.3	95.3	98.9	97.1
	$\frac{1}{2}\log(2)$	-26.5	-17.9	-16.9	0.092	0.082	0.243	0.289	0.280	0.489	92.2	94.1	99.4	97.3
	$\frac{1}{2}\log(3)$	-18.0	-10.8	-0.4	0.091	0.081	0.253	0.285	0.278	0.504	92.6	94.6	99.3	97.3
25%	$\frac{1}{2}\log(1.5)$	-26.5	-16.4	-5.9	0.020	0.019	0.046	0.131	0.132	0.213	93.5	93.7	98.6	95.1
	$\frac{1}{2}\log(2)$	-22.2	-14.2	0.1	0.023	0.018	0.043	0.130	0.126	0.206	91.8	94.2	98.8	95.9
	$\frac{1}{2}\log(3)$	-23.2	-14.7	-1.1	0.035	0.026	0.050	0.138	0.140	0.223	84.1	88.5	99.6	95.8
50%	$\frac{1}{2}\log(1.5)$	-27.8	-15.4	-4.6	0.016	0.013	0.032	0.112	0.111	0.178	92.5	93.8	98.4	94.8
	$\frac{1}{2}\log(2)$	-26.4	-17.9	-4.5	0.021	0.016	0.030	0.111	0.111	0.173	88.2	92.3	99.4	95.0
	$\frac{1}{2}\log(3)$	-25.4	-17.7	-3.5	0.032	0.022	0.031	0.113	0.111	0.176	77.5	84.6	99.3	95.4

Disease Prevalence	$\beta_{x,3}$	Percent Bias $\hat{\beta}^{(N)}_{x,3}$	Percent Bias $\hat{\beta}^{(C)}_{x,3}$	Percent Bias $\hat{\beta}^{(E)}_{x,3}$	MSE $\hat{\beta}^{(N)}_{x,3}$	MSE $\hat{\beta}^{(C)}_{x,3}$	MSE $\hat{\beta}^{(E)}_{x,3}$	SE $\hat{\beta}^{(N)}_{x,3}$	SE $\hat{\beta}^{(C)}_{x,3}$	SE $\hat{\beta}^{(E)}_{x,3}$	Coverage Rate $\hat{\beta}^{(N)}_{x,3}$	Coverage Rate $\hat{\beta}^{(C)}_{x,3}$	Coverage Rate $\hat{\beta}^{(E)}_{x,3}$	Coverage Rate $\hat{\beta}^{(E*)}_{x,3}$
5%	$\log(1.5)$	-21.6	-15.1	1.2	0.070	0.071	0.092	0.250	0.259	0.303	93.0	94.3	98.6	96.1
	$\log(2)$	-21.9	-16.1	1.0	0.087	0.082	0.100	0.253	0.264	0.316	89.7	92.0	98.3	95.8
	$\log(3)$	-23.9	-18.1	-0.5	0.131	0.107	0.124	0.249	0.261	0.352	78.7	87.6	98.5	96.6
25%	$\log(1.5)$	-25.0	-19.5	-4.4	0.025	0.023	0.022	0.122	0.129	0.147	86.9	90.6	98.0	94.5
	$\log(2)$	-23.7	-17.9	-1.9	0.041	0.031	0.022	0.120	0.127	0.149	75.0	84.6	97.9	94.3
	$\log(3)$	-24.8	-18.8	-2.5	0.090	0.061	0.028	0.128	0.137	0.164	41.8	62.8	98.5	93.6
50%	$\log(1.5)$	-23.4	-17.6	-2.4	0.021	0.019	0.017	0.109	0.116	0.130	84.5	89.2	98.3	95.2
	$\log(2)$	-24.4	-18.3	-3.0	0.041	0.029	0.018	0.110	0.112	0.132	64.3	79.3	98.4	94.1
	$\log(3)$	-24.5	-18.6	-3.2	0.084	0.055	0.021	0.109	0.116	0.140	30.8	54.5	98.2	92.7

NOTE: Percent bias and MSE were computed by averaging $(\hat{\beta} - \beta)/\beta$ and $(\beta - \hat{\beta})^2$ over 1000 simulations. Standard error (SE) is the square root of the empirical variance over all replicates. Coverage rate represents the coverage of a 95% confidence interval. We used the sandwich variances for the confidence intervals of the naive and cut-off calibration methods. For the exact calibration method, we applied both the sandwich and hessian matrix approaches for estimating the variances in the coverage rate evaluation, denoting by $\hat{\beta}^{(E)}$ and $\hat{\beta}^{(E*)}$ in the "Coverage Rate" columns respectively.

Table S8. *Comparison of operating characteristics in presence of negative variances for naive ($\hat{\beta}^{(N)}$), cut-off calibration ($\hat{\beta}^{(C)}$) and exact calibration ($\hat{\beta}^{(E)}$) methods, under Scenario II in Appendix A, i.e $\sigma_1^2 \sim U(0.25, 0.75)$ and $\sigma_d^2 \sim U(1.5, 2.5)$ for $d = 0, 2$.*

Disease		Percent Bias			MSE			SE			Coverage Rate			
Prevalence	$\beta_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(E*)}_{x,2}$
5%	$\frac{1}{2}\log(1.5)$	-22.7	-14.8	-15.8	0.075	0.072	0.106	0.271	0.266	0.324	94.9	95.5	100.0	95.3
	$\frac{1}{2}\log(2)$	-12.4	-9.9	-2.8	0.081	0.080	0.115	0.282	0.281	0.340	94.7	94.3	100.0	95.8
	$\frac{1}{2}\log(3)$	-13.7	-10.4	-5.7	0.088	0.087	0.123	0.287	0.290	0.349	94.0	93.4	99.9	95.2
25%	$\frac{1}{2}\log(1.5)$	-13.3	-8.3	-0.5	0.018	0.017	0.024	0.132	0.130	0.155	94.7	94.6	100.0	95.0
	$\frac{1}{2}\log(2)$	-16.0	-12.5	-5.7	0.020	0.018	0.023	0.129	0.125	0.151	93.4	95.2	100.0	96.5
	$\frac{1}{2}\log(3)$	-15.4	-12.0	-5.3	0.026	0.023	0.026	0.139	0.135	0.160	90.0	92.8	100.0	95.5
50%	$\frac{1}{2}\log(1.5)$	-15.4	-12.4	-4.1	0.014	0.013	0.018	0.114	0.111	0.135	93.0	94.1	100.0	93.7
	$\frac{1}{2}\log(2)$	-16.3	-13.4	-5.4	0.016	0.014	0.018	0.113	0.110	0.133	91.6	93.4	99.9	94.5
	$\frac{1}{2}\log(3)$	-18.4	-15.1	-7.9	0.024	0.020	0.021	0.115	0.114	0.137	85.8	88.7	99.9	92.2

Disease		Percent Bias			MSE			SE			Coverage Rate			
Prevalence	$\beta_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(E*)}_{x,3}$
5%	$\log(1.5)$	-17.3	-13.6	-7.3	0.074	0.074	0.083	0.263	0.267	0.287	93.3	94.2	99.9	93.8
	$\log(2)$	-17.8	-14.5	-7.0	0.082	0.080	0.081	0.258	0.264	0.281	91.8	92.4	99.8	94.4
	$\log(3)$	-16.3	-12.9	-5.3	0.102	0.095	0.092	0.265	0.275	0.298	86.8	89.8	100.0	93.9
25%	$\log(1.5)$	-17.5	-14.0	-6.7	0.023	0.022	0.021	0.133	0.136	0.144	89.2	91.0	100.0	93.4
	$\log(2)$	-17.7	-14.6	-7.6	0.032	0.027	0.022	0.129	0.131	0.140	83.3	86.9	100.0	92.3
	$\log(3)$	-18.7	-15.5	-8.3	0.060	0.047	0.029	0.132	0.133	0.143	62.3	73.2	100.0	89.1
50%	$\log(1.5)$	-17.1	-13.9	-6.9	0.017	0.015	0.015	0.109	0.111	0.118	90.3	92.7	100.0	94.2
	$\log(2)$	-18.8	-15.5	-8.5	0.029	0.024	0.018	0.110	0.113	0.120	76.9	82.8	99.9	90.4
	$\log(3)$	-18.0	-14.9	-7.6	0.052	0.040	0.022	0.112	0.113	0.122	55.7	69.5	100.0	90.1

NOTE: Percent bias and MSE were computed by averaging $(\hat{\beta} - \beta)/\beta$ and $(\beta - \hat{\beta})^2$ over 1000 simulations. Standard error (SE) is the square root of the empirical variance over all replicates. Coverage rate represents the coverage of a 95% confidence interval. We used the sandwich variances for the confidence intervals of the naive and cut-off calibration methods. For the exact calibration method, we applied both the sandwich and hessian matrix approaches for estimating the variances in the coverage rate evaluation, denoting by $\hat{\beta}^{(E)}$ and $\hat{\beta}^{(E*)}$ in the "Coverage Rate" columns respectively.

Table S9. *Comparison of operating characteristics in presence of negative variances for naive ($\hat\beta^{(N)}$), cut-off calibration ($\hat\beta^{(C)}$) and exact calibration ($\hat\beta^{(E)}$) methods, under Scenario III in Appendix A, i.e $\sigma_d^2 \sim U(0.25, 0.75)$ for $d = 1, 2, 3$.*

Disease		Percent Bias			MSE			SE			Coverage Rate			
Prevalence	$\beta_{x,2}$	$\hat\beta^{(N)}_{x,2}$	$\hat\beta^{(C)}_{x,2}$	$\hat\beta^{(E)}_{x,2}$	$\hat\beta^{(N)}_{x,2}$	$\hat\beta^{(C)}_{x,2}$	$\hat\beta^{(E)}_{x,2}$	$\hat\beta^{(N)}_{x,2}$	$\hat\beta^{(C)}_{x,2}$	$\hat\beta^{(E)}_{x,2}$	$\hat\beta^{(N)}_{x,2}$	$\hat\beta^{(C)}_{x,2}$	$\hat\beta^{(E)}_{x,2}$	$\hat\beta^{(E*)}_{x,2}$
5%	$\frac{1}{2}\log(1.5)$	-18.1	-15.9	-14.2	0.078	0.076	0.092	0.277	0.275	0.302	94.7	94.7	99.9	95.3
	$\frac{1}{2}\log(2)$	-2.7	-2.8	0.5	0.077	0.076	0.094	0.277	0.276	0.307	95.2	95.1	99.9	94.9
	$\frac{1}{2}\log(3)$	-9.7	-8.1	-6.7	0.089	0.090	0.109	0.294	0.297	0.329	94.4	94.6	99.9	95.1
25%	$\frac{1}{2}\log(1.5)$	-9.8	-8.7	-4.6	0.019	0.018	0.021	0.135	0.134	0.146	94.2	94.6	99.7	95.0
	$\frac{1}{2}\log(2)$	-10.0	-9.0	-6.2	0.019	0.018	0.022	0.134	0.132	0.146	94.5	94.6	99.9	95.4
	$\frac{1}{2}\log(3)$	-9.3	-8.4	-4.3	0.021	0.021	0.023	0.136	0.137	0.151	94.6	94.6	99.9	95.3
50%	$\frac{1}{2}\log(1.5)$	-7.7	-5.8	-2.0	0.013	0.013	0.015	0.111	0.112	0.121	95.4	94.7	99.8	94.6
	$\frac{1}{2}\log(2)$	-11.3	-10.5	-7.0	0.013	0.013	0.014	0.109	0.109	0.118	93.9	94.8	99.8	94.9
	$\frac{1}{2}\log(3)$	-11.1	-10.3	-5.8	0.016	0.015	0.016	0.110	0.109	0.121	92.4	93.1	99.8	94.2

Disease		Percent Bias			MSE			SE			Coverage Rate			
Prevalence	$\beta_{x,3}$	$\hat\beta^{(N)}_{x,3}$	$\hat\beta^{(C)}_{x,3}$	$\hat\beta^{(E)}_{x,3}$	$\hat\beta^{(N)}_{x,3}$	$\hat\beta^{(C)}_{x,3}$	$\hat\beta^{(E)}_{x,3}$	$\hat\beta^{(N)}_{x,3}$	$\hat\beta^{(C)}_{x,3}$	$\hat\beta^{(E)}_{x,3}$	$\hat\beta^{(N)}_{x,3}$	$\hat\beta^{(C)}_{x,3}$	$\hat\beta^{(E)}_{x,3}$	$\hat\beta^{(E*)}_{x,3}$
5%	log(1.5)	-9.8	-9.5	-5.8	0.069	0.070	0.075	0.261	0.263	0.273	94.7	94.7	99.9	94.6
	log(2)	-8.8	-8.2	-4.3	0.080	0.079	0.083	0.276	0.276	0.287	92.3	92.6	99.9	93.2
	log(3)	-10.4	-9.7	-5.5	0.088	0.088	0.089	0.274	0.277	0.292	92.2	92.2	99.9	93.4
25%	log(1.5)	-11.3	-10.5	-6.6	0.018	0.018	0.019	0.128	0.128	0.134	93.6	93.4	99.4	94.3
	log(2)	-11.2	-10.7	-6.8	0.022	0.021	0.019	0.125	0.125	0.131	91.6	91.8	99.9	93.1
	log(3)	-11.9	-11.4	-7.2	0.035	0.034	0.027	0.134	0.134	0.142	81.7	82.8	99.7	89.5
50%	log(1.5)	-10.6	-9.9	-6.0	0.014	0.014	0.014	0.111	0.112	0.115	93.5	93.9	100.0	95.2
	log(2)	-11.0	-10.6	-6.6	0.018	0.017	0.015	0.109	0.110	0.115	91.2	90.8	100.0	92.7
	log(3)	-11.9	-11.4	-7.4	0.030	0.028	0.021	0.111	0.112	0.118	78.7	79.2	100.0	88.3

NOTE: Percent bias and MSE were computed by averaging $(\hat\beta - \beta)/\beta$ and $(\beta - \hat\beta)^2$ over 1000 simulations. Standard error (SE) is the square root of the empirical variance over all replicates. Coverage rate represents the coverage of a 95% confidence interval. We used the sandwich variances for the confidence intervals of the naive and cut-off calibration methods. For the exact calibration method, we applied both the sandwich and hessian matrix approaches for estimating the variances in the coverage rate evaluation, denoting by $\hat\beta^{(E)}$ and $\hat\beta^{(E*)}$ in the "Coverage Rate" columns respectively.

Table S10

Comparison of operating characteristics over different sample sizes in both the contributed studies and calibration sub-study for naive ($\hat{\beta}^{(N)}$), cut-off calibration ($\hat{\beta}^{(C)}$) and exact calibration ($\hat{\beta}^{(E)}$) methods under the COCS design, disease prevalence 25%. Specifically, 5 studies were considered, with each 100, 200 or 500 individuals (10% subjects included in the calibration subset).

Sample size		Percent Bias			MSE			SE×100($\widehat{\text{SE}}$ × 100)			Coverage Rate			
per study	$\beta_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}(\hat{\beta}^{(E)}_{x,2},\hat{\beta}^{(E*)}_{x,2})$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(E*)}_{x,2}$
100	$\frac{1}{2}\log(1.2)$	-29.4	-4.2	5.2	0.073	0.072	0.200	26.8(26.4)	26.8(25.9)	44.7(47.8,42.6)	94.2	93.7	96.8	94.6
	$\frac{1}{2}\log(1.5)$	-30.1	-14.0	-1.8	0.072	0.067	0.179	26.2(26.7)	25.7(26.1)	42.3(48.2,42.9)	95.5	94.8	97.5	96.5
	$\frac{1}{2}\log(2)$	-27.1	-14.8	-6.6	0.080	0.076	0.201	26.7(27.1)	27.1(26.6)	44.8(52.0,44.3)	93.8	94.9	98.1	96.2
	$\frac{1}{2}\log(3)$	-20.1	-10.0	6.4	0.091	0.079	0.217	28.1(27.6)	27.6(27.3)	46.5(52.4,45.9)	92.1	93.7	97.9	96.8
200	$\frac{1}{2}\log(1.2)$	-33.0	-7.5	1.5	0.033	0.034	0.087	18.0(18.5)	18.4(18.1)	29.4(30.4,30.2)	95.7	95.3	97.0	96.9
	$\frac{1}{2}\log(1.5)$	-26.8	-11.7	0.9	0.036	0.033	0.091	18.3(18.7)	18.1(18.3)	30.1(31.8,30.6)	94.5	95.1	96.3	95.8
	$\frac{1}{2}\log(2)$	-31.1	-16.3	-3.9	0.050	0.041	0.100	19.5(19.0)	19.5(18.5)	31.6(32.6,31.3)	91.5	92.6	96.7	95.8
	$\frac{1}{2}\log(3)$	-22.6	-13.0	0.3	0.057	0.043	0.109	20.3(19.4)	19.4(19.0)	33.0(33.8,32.4)	88.6	92.8	95.5	94.8
500	$\frac{1}{2}\log(1.2)$	-34.7	-3.9	4.4	0.016	0.014	0.039	12.1(11.7)	11.7(11.4)	19.7(19.5,19.2)	92.8	94.1	95.4	95.0
	$\frac{1}{2}\log(1.5)$	-25.1	-8.7	6.1	0.018	0.014	0.040	12.4(11.8)	11.7(11.5)	20.0(19.8,19.5)	91.7	95.4	95.1	95.1
	$\frac{1}{2}\log(2)$	-25.4	-13.5	-0.7	0.022	0.015	0.038	11.8(11.9)	11.3(11.6)	19.5(20.3,19.9)	89.9	94.3	95.6	95.3
	$\frac{1}{2}\log(3)$	-24.5	-13.5	0.0	0.032	0.020	0.042	11.7(12.2)	11.9(11.9)	20.4(21.0,20.6)	79.6	88.7	95.9	95.5

Sample size		Percent Bias			MSE			SE×100($\widehat{\text{SE}}$ × 100)			Coverage Rate			
per study	$\beta_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}(\hat{\beta}^{(E)}_{x,3},\hat{\beta}^{(E*)}_{x,3})$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(E*)}_{x,3}$
100	$\frac{1}{2}\log(1.2)$	-17.8	-11.2	9.4	0.062	0.070	0.094	24.7(25.1)	26.4(26.5)	30.6(31.8,30.5)	96.1	95.3	97.2	96.6
	$\frac{1}{2}\log(1.5)$	-23.2	-16.5	0.5	0.076	0.077	0.096	26.0(25.1)	27.0(26.4)	31.1(31.9,30.5)	93.0	93.4	96.4	95.4
	$\frac{1}{2}\log(2)$	-22.1	-16.2	0.2	0.092	0.085	0.099	26.1(25.3)	26.9(26.7)	31.5(34.3,31.2)	89.4	92.5	96.7	96.0
	$\frac{1}{2}\log(3)$	-22.8	-16.2	1.4	0.129	0.108	0.113	25.7(25.7)	27.7(27.1)	33.6(35.7,32.7)	83.2	88.1	96.8	95.1
200	$\frac{1}{2}\log(1.2)$	-19.9	-12.6	3.1	0.033	0.036	0.048	17.9(17.6)	18.8(18.5)	21.8(21.8,21.3)	94.0	95.1	94.9	94.3
	$\frac{1}{2}\log(1.5)$	-23.1	-16.3	0.0	0.040	0.038	0.044	17.7(17.7)	18.4(18.5)	20.9(21.8,21.5)	91.9	93.4	95.5	95.1
	$\frac{1}{2}\log(2)$	-26.4	-19.7	-3.2	0.070	0.059	0.055	19.1(17.7)	20.0(18.6)	23.4(22.3,22.0)	79.1	87.7	94.3	93.8
	$\frac{1}{2}\log(3)$	-23.4	-17.0	0.6	0.100	0.073	0.055	18.3(18.0)	19.6(18.8)	23.5(23.5,22.7)	68.4	81.4	95.2	94.3
500	$\frac{1}{2}\log(1.5)$	-23.1	-15.7	0.7	0.014	0.015	0.018	11.2(11.1)	11.8(11.6)	13.6(13.5,13.4)	94.0	94.8	94.8	94.7
	$\frac{1}{2}\log(1.5)$	-23.6	-17.0	0.4	0.023	0.020	0.021	11.6(11.1)	12.2(11.7)	14.3(13.6,13.5)	84.9	89.8	94.6	94.2
	$\frac{1}{2}\log(1.5)$	-24.0	-17.2	0.4	0.040	0.028	0.019	11.1(11.2)	11.6(11.7)	13.7(13.9,13.7)	67.8	82.1	96.2	96.0
	$\frac{1}{2}\log(1.5)$	-24.8	-17.8	-0.1	0.087	0.053	0.022	11.3(11.3)	12.3(11.8)	14.7(14.6,14.3)	32.0	62.0	95.2	94.4

NOTE: Percent bias and MSE were computed by averaging $(\hat{\beta}-\beta)/\beta$ and $(\beta-\hat{\beta})^2$ over 1000 simulations. Empirical standard error (SE) is the square root of the empirical variance over all replicates. Coverage rate represents the coverage of a 95% confidence interval. We used the sandwich variances for the $\widehat{\text{SE}}$'s and confidence intervals of the naive and cut-off calibration methods. For the exact calibration method, we applied both the sandwich and hessian matrix approaches for estimating the variances in the coverage rate and estimated standard error evaluation, denoting by $\hat{\beta}^{(E)}$ and $\hat{\beta}^{(E*)}$ in the "Coverage Rate" columns respectively. In the SE×100($\widehat{\text{SE}} \times 100$) section, the numbers in the brackets denote the average estimated standard error ($\widehat{\text{SE}}$) of $\hat{\beta}$ over all replicates; for the exact calibration method, the first and second number denote the $\widehat{\text{SE}}$ by standard variance and pseudo-likelihood hessian matrix methods respectively.

TABLE S11

Comparison of operating characteristics for naive ($\hat{\beta}^{(N)}$), cut-off calibration ($\hat{\beta}^{(C)}$) and exact calibration ($\hat{\beta}^{(E)}$) methods under the COCS design with $\epsilon_{x_{jk}}$ following a normal,uniform and skew normal distribution, with disease prevalence 25%.

Distribution		Percent Bias			MSE			SE			Coverage Rate			
of $\epsilon_{x_{jk}}$	$\beta_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(N)}_{x,2}$	$\hat{\beta}^{(C)}_{x,2}$	$\hat{\beta}^{(E)}_{x,2}$	$\hat{\beta}^{(E*)}_{x,2}$
Normal	$\frac{1}{2}\log(1.2)$	-36.1	-5.6	4.9	0.008	0.007	0.019	0.081	0.081	0.137	92.7	95.4	96.0	95.9
	$\frac{1}{2}\log(1.5)$	-30.3	-13.2	-2.6	0.011	0.007	0.019	0.083	0.080	0.138	88.0	93.4	95.0	94.9
	$\frac{1}{2}\log(2)$	-26.1	-13.8	-0.4	0.015	0.009	0.019	0.085	0.084	0.138	80.6	89.8	96.2	95.8
	$\frac{1}{2}\log(3)$	-24.6	-13.1	-0.2	0.026	0.012	0.022	0.086	0.085	0.149	66.1	86.3	94.9	94.4
Uniform	$\frac{1}{2}\log(1.2)$	-35.4	-10.8	-3.4	0.008	0.006	0.018	0.083	0.079	0.134	93.3	95.8	95.5	95.4
	$\frac{1}{2}\log(1.5)$	-27.5	-11.7	1.2	0.010	0.007	0.019	0.084	0.082	0.137	89.9	93.6	95.8	95.4
	$\frac{1}{2}\log(2)$	-24.3	-12.3	1.6	0.014	0.008	0.018	0.084	0.081	0.135	83.2	92.3	96.2	96.2
	$\frac{1}{2}\log(3)$	-22.7	-12.5	0.8	0.022	0.012	0.021	0.083	0.084	0.143	70.3	88.1	95.3	95.2
Skew	$\frac{1}{2}\log(1.2)$	-42.7	-13.7	-11.7	0.009	0.007	0.019	0.084	0.081	0.138	92.2	95.3	95.7	95.6
	$\frac{1}{2}\log(1.5)$	-31.8	-13.0	-3.7	0.011	0.008	0.021	0.085	0.084	0.143	86.9	93.3	94.1	94.1
Normal	$\frac{1}{2}\log(2)$	-28.5	-14.7	-4.5	0.017	0.009	0.019	0.085	0.082	0.138	77.8	90.0	96.4	96.4
	$\frac{1}{2}\log(3)$	-27.3	-13.5	-1.8	0.030	0.013	0.023	0.085	0.087	0.150	57.0	84.4	95.5	95.4

Distribution		Percent Bias			MSE			SE			Coverage Rate			
of $\epsilon_{x_{jk}}$	$\beta_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(N)}_{x,3}$	$\hat{\beta}^{(C)}_{x,3}$	$\hat{\beta}^{(E)}_{x,3}$	$\hat{\beta}^{(E*)}_{x,3}$
Normal	log(1.2)	-20.8	-14.3	3.8	0.007	0.007	0.009	0.077	0.081	0.094	93.0	93.5	95.3	95.3
	log(1.5)	-24.1	-16.9	0.1	0.016	0.012	0.009	0.079	0.084	0.097	77.8	86.2	93.7	93.6
	log(2)	-24.7	-17.9	-0.6	0.037	0.023	0.011	0.086	0.090	0.105	42.4	67.2	93.4	93.1
	log(3)	-24.7	-17.5	0.1	0.080	0.044	0.011	0.080	0.086	0.104	8.1	37.1	94.7	94.0
Uniform	log(1.2)	-21.4	-16.2	1.3	0.008	0.008	0.009	0.080	0.082	0.094	92.1	93.5	94.6	94.5
	log(1.5)	-22.8	-16.4	0.4	0.015	0.011	0.010	0.081	0.084	0.098	77.0	86.3	94.4	94.2
	log(2)	-23.0	-16.5	0.8	0.031	0.020	0.009	0.077	0.081	0.095	49.2	73.8	94.5	94.5
	log(3)	-23.3	-17.0	0.9	0.072	0.042	0.010	0.079	0.084	0.101	11.5	38.3	96.0	96.0
Skew	log(1.2)	-20.1	-13.1	5.1	0.006	0.005	0.007	0.066	0.070	0.080	92.9	94.5	95.4	95.4
	log(1.5)	-23.1	-15.4	1.9	0.014	0.010	0.008	0.071	0.075	0.088	72.2	85.5	93.2	93.2
Normal	log(2)	-23.2	-16.1	1.9	0.031	0.018	0.008	0.070	0.074	0.087	34.6	64.5	94.9	94.8
	log(3)	-24.1	-16.7	1.8	0.075	0.039	0.008	0.069	0.074	0.088	2.7	31.1	94.4	94.4

NOTE: Percent bias and MSE were computed by averaging $(\hat{\beta} - \beta)/\beta$ and $(\beta - \hat{\beta})^2$ over 1000 simulations. Standard error (SE) is the square root of the empirical variance over all replicates. Coverage rate represents the coverage of a 95% confidence interval. We used the sandwich variances for the confidence intervals of the naive and cut-off calibration methods. For the exact calibration method, we applied both the sandwich and hessian matrix approaches for estimating the variances in the coverage rate evaluation, denoting by $\hat{\beta}^{(E)}$ and $\hat{\beta}^{(E*)}$ in the "Coverage Rate" columns respectively.

Table S12

Comparison of operating characteristics for naive ($\hat{\beta}^{(N)}$), cut-off calibration ($\hat{\beta}^{(C)}$) and exact calibration ($\hat{\beta}^{(E)}$) methods under the COCS design with $log(X_{jk}) \sim N(\alpha_{0j}, \sigma_x^2)$, with $\alpha_{0j} \sim N(4.5, 0.1^2)$ and $\sigma_x^2 = 0.01$, $j = 1, \ldots, 5$.

Disease Prevalence	$\beta_{x,2}$	Percent Bias $\hat{\beta}^{(N)}_{x,2}$	Percent Bias $\hat{\beta}^{(C)}_{x,2}$	Percent Bias $\hat{\beta}^{(E)}_{x,2}$	MSE $\hat{\beta}^{(N)}_{x,2}$	MSE $\hat{\beta}^{(C)}_{x,2}$	MSE $\hat{\beta}^{(E)}_{x,2}$	SE×100($\widehat{\text{SE}}\times 100$) $\hat{\beta}^{(N)}_{x,2}$	SE×100($\widehat{\text{SE}}\times 100$) $\hat{\beta}^{(C)}_{x,2}$	SE×100($\widehat{\text{SE}}\times 100$) $\hat{\beta}^{(E)}_{x,2}(\hat{\beta}^{(E)}_{x,2},\hat{\beta}^{(E*)}_{x,2})$	Coverage Rate $\hat{\beta}^{(N)}_{x,2}$	Coverage Rate $\hat{\beta}^{(C)}_{x,2}$	Coverage Rate $\hat{\beta}^{(E)}_{x,2}$	Coverage Rate $\hat{\beta}^{(E*)}_{x,2}$
5%	$\frac{1}{2}\log(1.2)$	-48.7	-6.5	-0.8	0.035	0.042	0.136	18.2(17.6)	20.4(19.4)	36.9(37.2,36.8)	94.0	94.7	96.7	96.2
	$\frac{1}{2}\log(1.5)$	-49.3	-24.0	-12.0	0.042	0.043	0.161	17.8(17.7)	20.2(19.3)	40.0(38.3,38.2)	91.2	94.2	96.2	95.9
	$\frac{1}{2}\log(2)$	-45.0	-21.0	8.2	0.058	0.051	0.282	18.3(18.1)	21.5(20.3)	53.1(52.5,51.5)	84.7	93.5	95.1	95.0
	$\frac{1}{2}\log(3)$	-44.8	-23.0	0.9	0.094	0.062	0.288	18.2(18.2)	21.5(20.6)	53.7(54.0,53.5)	71.3	88.4	96.5	96.0
25%	$\frac{1}{2}\log(1.2)$	-62.0	-24.4	-14.5	0.011	0.011	0.032	8.7(8.9)	10.3(10.4)	17.7(18.3,17.9)	90.7	93.1	96.4	95.9
	$\frac{1}{2}\log(1.5)$	-48.5	-22.1	2.1	0.017	0.012	0.032	8.7(8.9)	10.0(10.1)	18.0(18.4,17.9)	81.3	93.8	96.7	96.1
	$\frac{1}{2}\log(2)$	-46.4	-24.7	2.2	0.034	0.017	0.034	9.0(8.9)	10.0(10.5)	18.4(18.4,17.9)	56.2	86.8	95.6	94.8
	$\frac{1}{2}\log(3)$	-46.0	-25.4	1.0	0.073	0.031	0.041	9.3(9.1)	10.6(9.9)	20.3(20.3,19.1)	20.2	71.9	95.7	93.5
50%	$\frac{1}{2}\log(1.2)$	-58.8	-11.8	2.1	0.009	0.008	0.021	7.6(7.6)	8.7(8.8)	14.3(15.6,15.0)	89.5	95.4	97.2	96.4
	$\frac{1}{2}\log(1.5)$	-54.1	-22.2	-2.6	0.018	0.010	0.025	7.8(7.6)	9.1(9.1)	15.7(15.6,15.0)	69.5	92.0	94.8	93.7
	$\frac{1}{2}\log(2)$	-48.9	-24.8	-0.0	0.035	0.015	0.023	8.0(7.6)	8.7(8.4)	15.2(15.7,15.0)	38.9	82.4	96.4	95.3
	$\frac{1}{2}\log(3)$	-47.6	-27.5	0.6	0.074	0.031	0.026	7.8(7.6)	8.8(8.3)	16.2(16.1,15.0)	8.4	58.1	95.7	93.1

Disease Prevalence	$\beta_{x,3}$	Percent Bias $\hat{\beta}^{(N)}_{x,3}$	Percent Bias $\hat{\beta}^{(C)}_{x,3}$	Percent Bias $\hat{\beta}^{(E)}_{x,3}$	MSE $\hat{\beta}^{(N)}_{x,3}$	MSE $\hat{\beta}^{(C)}_{x,3}$	MSE $\hat{\beta}^{(E)}_{x,3}$	SE×100($\widehat{\text{SE}}\times 100$) $\hat{\beta}^{(N)}_{x,3}$	SE×100($\widehat{\text{SE}}\times 100$) $\hat{\beta}^{(C)}_{x,3}$	SE×100($\widehat{\text{SE}}\times 100$) $\hat{\beta}^{(E)}_{x,3}(\hat{\beta}^{(E)}_{x,3},\hat{\beta}^{(E*)}_{x,3})$	Coverage Rate $\hat{\beta}^{(N)}_{x,3}$	Coverage Rate $\hat{\beta}^{(C)}_{x,3}$	Coverage Rate $\hat{\beta}^{(E)}_{x,3}$	Coverage Rate $\hat{\beta}^{(E*)}_{x,3}$
5%	$\frac{1}{2}\log(1.2)$	-40.7	-29.8	0.3	0.043	0.075	0.127	19.3(19.1)	26.8(26.3)	35.6(33.9,33.5)	92.3	94.5	94.8	94.0
	$\frac{1}{2}\log(1.5)$	-39.8	-26.3	6.5	0.062	0.074	0.107	19.0(19.0)	25.1(26.3)	32.6(33.9,33.5)	86.1	92.5	95.3	95.1
	$\frac{1}{2}\log(2)$	-41.0	-28.0	6.0	0.115	0.101	0.222	18.4(19.0)	25.3(25.4)	47.0 (49.8,48.1)	65.9	89.1	95.7	95.5
	$\frac{1}{2}\log(3)$	-40.8	-28.4	4.3	0.238	0.172	0.239	19.4(19.0)	27.4(27.5)	48.7(48.5,45.1)	36.2	75.2	95.1	94.2
25%	$\frac{1}{2}\log(1.2)$	-41.2	-31.0	0.0	0.015	0.021	0.028	9.5(9.7)	13.5(13.8)	16.6(16.3,16.0)	87.4	92.9	95.5	95.2
	$\frac{1}{2}\log(1.5)$	-40.9	-30.2	0.2	0.037	0.032	0.028	9.8(9.6)	13.1(13.8)	16.6(16.4,16.0)	58.2	83.9	95.3	94.2
	$\frac{1}{2}\log(2)$	-40.6	-29.0	1.8	0.088	0.057	0.032	9.6(9.6)	12.9(12.7)	17.7(17.3,16.4)	17.3	64.8	94.5	93.7
	$\frac{1}{2}\log(3)$	-41.5	-29.2	2.0	0.218	0.122	0.040	10.3(9.7)	13.6(12.1)	20.0(19.0,17.0)	0.7	36.3	94.9	91.1
50%	$\frac{1}{2}\log(1.2)$	-38.6	-29.4	1.2	0.012	0.015	0.017	8.2(8.3)	11.1(11.4)	13.2(13.9,13.7)	86.8	92.3	96.6	96.3
	$\frac{1}{2}\log(1.5)$	-40.9	-29.9	-1.1	0.035	0.028	0.020	8.5(8.3)	11.4(11.8)	14.1(14.1,13.7)	47.2	81.4	95.2	94.4
	$\frac{1}{2}\log(2)$	-40.1	-28.4	1.0	0.085	0.051	0.021	8.8(8.3)	11.0(11.9)	14.6(15.0,13.9)	10.4	60.0	96.0	94.0
	$\frac{1}{2}\log(3)$	-41.2	-29.9	0.1	0.213	0.188	0.034	8.4(8.4)	11.3(11.9)	14.7(13.8,12.7)	0.2	27.9	93.7	89.0

NOTE: Percent bias and MSE were computed by averaging $(\hat{\beta}-\beta)/\beta$ and $(\beta-\hat{\beta})^2$ over 1000 simulations. Empirical standard error (SE) is the square root of the empirical variance over all replicates. Coverage rate represents the coverage of a 95% confidence interval. We used the sandwich variances for the $\widehat{\text{SE}}$'s and confidence intervals of the naive and cut-off calibration methods. For the exact calibration method, we applied both the sandwich and hessian matrix approaches for estimating the variances in the coverage rate and estimated standard error evaluation, denoting by $\hat{\beta}^{(E)}$ and $\hat{\beta}^{(E*)}$ in the "Coverage Rate" columns respectively. In the SE×100($\widehat{\text{SE}}\times 100$) section, the numbers in the brackets denote the average estimated standard error ($\widehat{\text{SE}}$) of $\hat{\beta}$ over all replicates; for the exact calibration method, the first and second number denote the $\widehat{\text{SE}}$ by standard variance and pseudo-likelihood hessian matrix methods respectively.

REFERENCES

Sloan, A., Song, Y., Gail, M. H., Betensky, R., Rosner, B., Ziegler, R. G., Smith-Warner, S. A. and Wang, M. (2019). Design and analysis considerations for combining data from multiple biomarker studies. *Stat. Med.* **38** 1303–1320.

C. Cheng
Center for Methods in Implementation and Prevention Science
Yale School of Public Health
135 College Street
New Haven, Connecticut 06511
and
Department of Mathematical Sciences
Tsinghua University
Beijing, China
E-mail: c.cheng@yale.edu

B. Rosner
Channing Laboratory
180 Longwood Avenue
Boston, Massachusetts 02115
and
Department of Biostatistics
Harvard T.H. Chan School of Public Health
Kresge Building
Boston, Massachusetts 02115
E-mail: stbar@channing.harvard.edu

M. Wang
Channing Division of Network Medicine
Brigham and Women's Hospital
Harvard Medical School
180 Longwood Avenue
Boston, Massachusetts 02115
and
Department of Epidemiology and Biostatistics
Harvard T.H. Chan School of Public Health
Kresge Building
Boston, Massachusetts 02115
E-mail: stmow@channing.harvard.edu