BIOMETRICS 63, 1–5 DOI: xx.xxxx/j.xxxx-xxxx.xxxx.xxxxx.x
December 2007

Supplementary Material: Statistical Methods for Analysis of Combined Biomarker Data from Multiple Nested Case-Control Studies*

Chao Cheng

Center for Methods in Implementation and Prevention Science, Yale School of Public Health, New Haven

**email:* c.cheng@yale.edu

and

Abigail Sloan

Department of Biostatistics, Harvard School of Public Health, Boston

**email:* abigail.sloan@pfizer.com

and

Molin Wang

Departments of Epidemiology and Biostatistics, Harvard School of Public Health, Boston

**email:* stmow@channing.harvard.edu

*Note: This is a working paper, and the authorship line may change

Appendix

Appendix A: Proof for the Conditions for Assumption (6)

In this section, we show that under under either of the three conditions, we have $f(X_{jkm}|H_{jkm,d}, \boldsymbol{W}_{jkm}, Y_{jkm}) \approx f(X_{jkm}|H_{jkm,d}, \boldsymbol{W}_{jkm})$.

The first condition is that σ_d^2 is small for $d = 1, \dots, J$. Specifically, if $\sigma_d^2 \to 0$, then $\epsilon_{jkm,d} \to 0$, which implies $X_{jkm} \to \frac{H_{jkm,d}-\xi_d}{1+\gamma_d}$. It means both $f(X_{jkm}|H_{jkm,d}, \boldsymbol{W}_{jkm}, Y_{jkm})$ and $f(X_{jkm}|H_{jkm,d}, \boldsymbol{W}_{jkm})$ converge to a degenerate distribution, i.e., $\mathbb{I}(X_{jkm} = \frac{H_{jkm,d}-\xi_d}{1+\gamma_d})$. This completes the proof for the first condition.

Now, we consider the second and third conditions, i.e., small exposure effect and rare disease prevalence. In the beginning, we can show that (some subscripts are suppressed for brevity):

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

where $\boldsymbol{Z}^*$ denotes the common variables of $\boldsymbol{Z}$ and $\boldsymbol{W}$. If the exposure effect is small ($\beta_x \approx 0$), Y is nearly independent with X, i.e $f(Y|X, \boldsymbol{Z}^*) \approx f(Y|\boldsymbol{Z}^*)$, which implies

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

This finishes proof for the second condition. If the disease prevalence is rare, then $f(Y|X, \boldsymbol{Z}^*) \approx$

$f(Y=0|X,\boldsymbol{Z}^*) \approx 1$, which leads to

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

This completes the proof for the third condition.

Appendix B: Monte Carlo and GHQ Approaches

Now, we introduce the Monte Carlo and GHQ approaches respectively.

- Monte Carlo approach: Draw G samples $\boldsymbol{X}_{jk}^{(1)},\ldots,\boldsymbol{X}_{jk}^{(G)}$ i.i.d. from $N(\hat{\boldsymbol{\mu}}_{jk};\hat{\boldsymbol{s}}_{jk}^2)$, where $\boldsymbol{X}_{jk}^{(g)} = [\boldsymbol{X}_{jk1}^{(g)},\ldots,\boldsymbol{X}_{jkM_{jk}}^{(g)}]^T$, $g=1,2,\ldots,G$. The likelihood contribution (18) can be numerically calculated by

$$
\tilde{L}_{jk}^{(E1)} = \frac{1}{G}\sum_{g=1}^{G}\frac{\exp\left\{\sum_{m=1}^{M_{jk}^{(1)}}(\beta_x X_{jkm}^{(g)} + \boldsymbol{\beta}_z^T\boldsymbol{Z}_{jkm})\right\}}{\sum_{\mathcal{M}\in\mathcal{C}_{jk}}\exp\left\{\sum_{m\in\mathcal{M}}(\beta_x X_{jkm}^{(g)} + \boldsymbol{\beta}_z^T\boldsymbol{Z}_{jkm})\right\}}.
$$

- Gauss-Hermite Quadrature (GHQ) approach: To calculate the integration to a function with respect to a multivariate normal distribution, the GHQ approach arranges P knots and their corresponding weights to each dimension (a total of P^D knots if there were D dimensions). Due to the knot numbers of the GHQ approach is sensitive to the dimension of the integration , we implement an integration dimension reduction strategy first before applying the GHQ approach to calculate (18). Specifically, let $B_{jkm} = X_{jkm} - X_{jk1}$, where $m=2,\ldots,M_{jk}$, and $B_{jk1}=0$. The conditional likelihood in (**??**) can be transformed to

$$
\tilde{L}_{jk} = \int \frac{\exp\left\{\sum_{m=1}^{M_{jk}^{(1)}}\left(\beta_x B_{jkm} + \beta_z^T(\boldsymbol{Z}_{jkm} - \boldsymbol{Z}_{jk1})\right)\right\}}{\sum_{\mathcal{M}\in\mathcal{C}_{jk}}\exp\left\{\sum_{m\in\mathcal{M}}\left(\beta_x B_{jkm} + \beta_z^T(\boldsymbol{Z}_{jkm} - \boldsymbol{Z}_{jk1})\right)\right\}} f\left(\boldsymbol{B}_{jk}|\tilde{\boldsymbol{\mu}}_{jk};\tilde{\boldsymbol{s}}_{jk}^2\right)\mathrm{d}\boldsymbol{B}_{jk}
$$

  where $\boldsymbol{B}_{jk} = [B_{jk2},\ldots,B_{jkM_{jk}}]$ follows a $(M_{jk}-1)$-dimension normal distribution with mean $\tilde{\boldsymbol{\mu}}_{jk} = [\hat{\mu}_{jk2} - \hat{\mu}_{jk1},\ldots,\hat{\mu}_{jkM_{jk}} - \hat{\mu}_{jk1}]^T$, and the variance matrix $\tilde{\boldsymbol{s}}_{jk}^2$ a $(M_{jk}-1)\times(M_{jk}-1)$ matrix with diagonal $(s_{jk2}^2 + s_{jk1}^2,\ldots,s_{jkM_{jk}}^2 + s_{jk1}^2)$ and all other elements s_{jk1}^2. We use this transformation to reduce the integration dimension from M_{jk} to $M_{jk}-1$. Now, based on the GHQ, the likelihood contribution $\tilde{L}_{jk}$ can be numerically calculated by

$$\tilde{L}_{jk}^{(E2)} = \sum_{p=1}^{P^{M_{jk}-1}} \frac{\omega_{jk}^{(p)*} \exp\left\{ \sum_{m=1}^{M_{jk}^{(1)}} \left(\beta_x \tilde{B}_{jkm}^{(p)} + \beta_z^T(\boldsymbol{Z}_{jkm} - \boldsymbol{Z}_{jk1})\right)\right\}}{\sum_{\mathcal{M}\in\mathcal{C}_{jk}} \exp\left\{ \sum_{m\in\mathcal{M}} \left(\beta_x \tilde{B}_{jkm}^{(p)} + \beta_z^T(\boldsymbol{Z}_{jkm} - \boldsymbol{Z}_{jk1})\right)\right\}}$$

where each dimension contains P knots (a total of $P^{M_{jk}-1}$ knots as there are totally $M_{jk}-1$ dimensions), $\tilde{\boldsymbol{B}}_{jk}^{(p)} = [\tilde{B}_{jk2}^{(p)}, \ldots, \tilde{B}_{jkM_{jk}}^{(p)}]$ terms ($p = 1, \ldots, P^{M_{jk}-1}$) are the knots, $\omega_{jk}^{(p)*}$ terms are the weights, and $B_{jk1}^{(p)}$ is fixed as zero for all $p = 1, \ldots, P^{M_{jk}-1}$. The weights, $\boldsymbol{\omega}_{jk}^{(p)*}$, and the knots $\tilde{\boldsymbol{B}}_{jk}^{(p)}$ ($p = 1, \ldots, P^{M_{jk}-1}$) are calculated based on P, $M_{jk}-1$, $\tilde{\boldsymbol{\mu}}_{jk}$ and $\tilde{\boldsymbol{s}}_{jk}^2$. For more details about the calculation of knots and weights, please see Jackel (2005). Moreover, R package "MultiGHQuad" can be applied to calculate the weights and knots.

Estimates of $\boldsymbol{\beta}$ can be obtained by maximizing the pseudo-likelihood $\tilde{L}^{(E1)} = \Pi_{j,k}\tilde{L}_{jk}^{(E1)}$ or $\tilde{L}^{(E2)} = \Pi_{j,k}\tilde{L}_{jk}^{(E2)}$. In the simulation studies and real data example, we choose G and P such that $G = 50$ and $P = 5$.

Appendix C: Approximate Likelihood Derivation in the Approximate Calibration Method

From equation (8), we learn the likelihood contribution from the k^{th} matched set in the j^{th} study is

$$L_{jk} \approx E_{\boldsymbol{X}_{jk}|\boldsymbol{H}_{jk},\boldsymbol{W}_{jk}}\left[\frac{\exp\left\{ \sum_{m=1}^{M_{jk}^{(1)}}(\beta_x X_{jkm} + \beta_z^T \boldsymbol{Z}_{jkm})\right\}}{\sum_{\mathcal{M}\in\mathcal{C}_{jk}} \exp\left\{ \sum_{m\in\mathcal{M}}(\beta_x X_{jkm} + \beta_z^T \boldsymbol{Z}_{jkm})\right\}}\right].$$

where $\boldsymbol{X}_{jk}|\boldsymbol{H}_{jk}, \boldsymbol{W}_{jk} \sim N(\boldsymbol{\mu}_{jk}, \boldsymbol{s}_{jk})$. In this section, we prove that under certain conditions, the likelihood contribution above can be further approximated by

$$L_{jk}^{(A)} = \frac{\exp\left\{ \sum_{m=1}^{M_{jk}^{(1)}}(\beta_x \mu_{jkm} + \beta_z^T \boldsymbol{Z}_{jkm})\right\}}{\sum_{\mathcal{M}\in\mathcal{C}_{jk}} \exp\left\{ \sum_{m\in\mathcal{M}}(\beta_x \mu_{jkm} + \beta_z^T \boldsymbol{Z}_{jkm})\right\}} \tag{1}$$

which is the likelihood contribution used by the approximate calibration method. Specifically, denoting the expression inside the expectation (8) by G, this likelihood contribution can be approximated by replacing G with a second order Taylor series expansion with respect to

$\boldsymbol{X}_{jk}$ about its expectation $\boldsymbol{\mu}_{jk}$ such that

$$\begin{aligned} G \approx & G|_{\boldsymbol{X}_{jk}=\boldsymbol{\mu}_{jk}} + \sum_{m=1}^{M_{jk}} \frac{\partial G}{\partial X_{jkm}}\Big|_{\boldsymbol{X}_{jk}=\boldsymbol{\mu}_{jk}} \Delta_{jkm} + \sum_{m=1}^{M_{jk}} \frac{\partial^2 G}{\partial X_{jkm}^2}\Big|_{\boldsymbol{X}_{jk}=\boldsymbol{\mu}_{jk}} \Delta_{jkm}^2 \\ & + 2 \sum_{1 \leqslant m < m' \leqslant M_{jk}} \frac{\partial^2 G}{\partial X_{jkm} X_{jkm'}} \Delta_{jkm} \Delta_{jkm'} + \text{remander}, \end{aligned} \tag{2}$$

where $\Delta_{jkm} = X_{jkm} - \mu_{jkm}$. Replacing this expansion in the expectation causes $\frac{\partial G}{\partial X_{jkm}}\Big|_{\boldsymbol{X}_{jk}=\boldsymbol{\mu}_{jk}} \Delta_{jkm}$ and $\frac{\partial^2 G}{\partial X_{jkm} X_{jkm'}} \Delta_{jkm} \Delta_{jkm'}$ to appear, as $E_{\boldsymbol{X}_{jk}|\boldsymbol{H}_{jk},\boldsymbol{W}_{jk}} \Delta_{jkm} = 0$ and $E_{\boldsymbol{X}_{jk}|\boldsymbol{H}_{jk},\boldsymbol{W}_{jk}} \Delta_{jkm} \Delta_{jkm'} = 0$ (Notice X_{jkm} is independent with $X_{jkm'}$). The second order term with respect to Δ_{jkm}^2 can be rewritten as a function of the $\mathrm{Var}(X_{jkm}|\boldsymbol{H}_{jkm}, \boldsymbol{W}_{jkm})$ and β_x such that

$$L_{jk} \approx G|_{\boldsymbol{X}_{jk}=\boldsymbol{\mu}_{jk}} + \sum_{m=1}^{M_{jk}} \frac{\partial^2 G}{\partial X_{jkm}^2}\Big|_{\boldsymbol{X}_{jk}=\boldsymbol{\mu}_{jk}} \mathrm{Var}(X_{jkm}|\boldsymbol{H}_{jkm}, \boldsymbol{W}_{jkm}), \tag{3}$$

where

$$\frac{\partial^2 G}{\partial X_{jkm}^2} = \beta_x^2 G \Big\{ \mathbb{I}_{(m \leqslant M_{jk}^{(1)})} (1 - \frac{3 \sum_{\mathcal{M} \in \mathcal{C}_{jk}} \mathbb{I}_{(\mathrm{m} \in \mathcal{M})} \exp\{V_{jkm}\}}{\sum_{\mathcal{M} \in \mathcal{C}_{jk}} \exp\{V_{jkm}\}}) + \frac{2 (\sum_{\mathcal{M} \in \mathcal{C}_{jk}} \mathbb{I}_{(\mathrm{m} \in \mathcal{M})} \exp\{V_{jkm}\})^2}{(\sum_{\mathcal{M} \in \mathcal{C}_{jk}} \exp\{V_{jkm}\})^2} \Big\},$$

and $V_{jkm} = \beta_x X_{jkm} + \beta_z^T \boldsymbol{Z}_{jkm}$. If $|\beta_x|$ or $\mathrm{Var}(X_{jkm}|\boldsymbol{H}_{jkm}, \boldsymbol{W}_{jkm})$ (i.e., σ_d^2 for $d = 0, \ldots, M$) is small, the second-order term also approaches zero. Therefore, for matched set k in study j, the likelihood contribution can be approximated by

$$L_{jk} \approx \frac{\exp\big\{ \sum_{m=1}^{M_{jk}^{(1)}} (\beta_x \mu_{jkm} + \beta_z^T \boldsymbol{Z}_{jkm}) \big\}}{\sum_{\mathcal{M} \in \mathcal{C}_{jk}} \exp\big\{ \sum_{m \in \mathcal{M}} (\beta_x \mu_{jkm} + \beta_z^T \boldsymbol{Z}_{jkm}) \big\}}.$$

This completes the proof.

Appendix D: Variance estimates for $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{r}}$

The point estimates for fixed effects $\boldsymbol{\theta}$ and the slope and intercept biases $\boldsymbol{r}$ are

$$\hat{\boldsymbol{\theta}} = \big(\boldsymbol{U}^T \widehat{\boldsymbol{V}}^{-1} \boldsymbol{U}\big)^{-1} \boldsymbol{U}^T \widehat{\boldsymbol{V}}^{-1} \boldsymbol{H},$$

$$\hat{\boldsymbol{r}} = \widehat{\boldsymbol{R}} \widehat{\boldsymbol{D}}^T \widehat{\boldsymbol{V}}^{-1} \big(\boldsymbol{H} - \boldsymbol{U} \hat{\boldsymbol{\theta}}\big),$$

where $\widehat{\boldsymbol{V}}$ and $\widehat{\boldsymbol{D}}$ are the abbreviations of $\mathcal{V}(\hat{\boldsymbol{\sigma}}^2, \hat{\boldsymbol{\theta}}^*, \hat{\boldsymbol{r}}^*)$ and $\widehat{\boldsymbol{D}} = \boldsymbol{\mathcal{D}}(\hat{\boldsymbol{\theta}})$ respectively. Here, $\hat{\boldsymbol{\theta}}^*$ and $\hat{\boldsymbol{r}}^*$ represent the point estimators of $\boldsymbol{\theta}$ and $\boldsymbol{r}$ in the second-to-the-last iteration. We can calculate $\widehat{\mathrm{Var}}(\hat{\boldsymbol{\theta}})$ and $\widehat{\mathrm{Var}}(\hat{\boldsymbol{r}})$ with fixed $\widehat{\boldsymbol{V}}$ and $\widehat{\boldsymbol{D}}$. In paricular, we have

$$\begin{aligned}
\widehat{\mathrm{Var}}(\hat{\boldsymbol{\theta}}) &= \mathrm{Var}\Big(\big(\boldsymbol{U}^T\widehat{\boldsymbol{V}}^{-1}\boldsymbol{U}\big)^{-1}\boldsymbol{U}^T\widehat{\boldsymbol{V}}^{-1}\boldsymbol{H}\Big)\\
&= \big(\boldsymbol{U}^T\widehat{\boldsymbol{V}}^{-1}\boldsymbol{U}\big)^{-1}\boldsymbol{U}^T\widehat{\boldsymbol{V}}^{-1}\mathrm{Var}(\boldsymbol{H})\Big(\big(\boldsymbol{U}^T\widehat{\boldsymbol{V}}^{-1}\boldsymbol{U}\big)^{-1}\boldsymbol{U}^T\widehat{\boldsymbol{V}}^{-1}\Big)^T\\
&= \big(\boldsymbol{U}^T\widehat{\boldsymbol{V}}^{-1}\boldsymbol{U}\big)^{-1}\boldsymbol{U}^T\widehat{\boldsymbol{V}}^{-1}\widehat{\boldsymbol{V}}\big(\boldsymbol{U}^T\widehat{\boldsymbol{V}}^{-1}\big)^T\big(\boldsymbol{U}^T\widehat{\boldsymbol{V}}^{-1}\boldsymbol{U}\big)^{-1}\\
&= \big(\boldsymbol{U}^T\widehat{\boldsymbol{V}}^{-1}\boldsymbol{U}\big)^{-1},
\end{aligned}\tag{4}$$

and

$$\begin{aligned}
\widehat{\mathrm{Var}}(\hat{\boldsymbol{r}}) &= \mathrm{Var}\Big(\widehat{\boldsymbol{R}}\widehat{\boldsymbol{D}}^T\widehat{\boldsymbol{V}}^{-1}\big(\boldsymbol{H}-\boldsymbol{U}\hat{\boldsymbol{\theta}}\big)\Big)\\
&= \widehat{\boldsymbol{R}}\widehat{\boldsymbol{D}}^T\widehat{\boldsymbol{V}}^{-1}\mathrm{Var}\Big(\boldsymbol{H}-\boldsymbol{U}\hat{\boldsymbol{\theta}}\Big)\Big(\widehat{\boldsymbol{R}}\widehat{\boldsymbol{D}}^T\widehat{\boldsymbol{V}}^{-1}\Big)^T\\
&= \widehat{\boldsymbol{R}}\widehat{\boldsymbol{D}}^T\widehat{\boldsymbol{V}}^{-1}\Big(\widehat{\boldsymbol{V}}-\boldsymbol{U}\mathrm{Var}(\hat{\boldsymbol{\theta}})\boldsymbol{U}^T\Big)\widehat{\boldsymbol{V}}^{-1}\widehat{\boldsymbol{D}}\widehat{\boldsymbol{R}}\\
&= \widehat{\boldsymbol{R}}\widehat{\boldsymbol{D}}^T\widehat{\boldsymbol{V}}^{-1}\Big(\widehat{\boldsymbol{V}}-\boldsymbol{U}\big(\boldsymbol{U}^T\widehat{\boldsymbol{V}}^{-1}\boldsymbol{U}\big)^{-1}\boldsymbol{U}^T\Big)\widehat{\boldsymbol{V}}^{-1}\widehat{\boldsymbol{D}}\widehat{\boldsymbol{R}}\\
&= \widehat{\boldsymbol{R}}\widehat{\boldsymbol{D}}^T\big(\widehat{\boldsymbol{V}}^{-1}-\widehat{\boldsymbol{V}}^{-1}\boldsymbol{U}(\boldsymbol{U}^T\widehat{\boldsymbol{V}}^{-1}\boldsymbol{U})\boldsymbol{U}^T\widehat{\boldsymbol{V}}^{-1}\big)\widehat{\boldsymbol{D}}\widehat{\boldsymbol{R}}.
\end{aligned}\tag{5}$$

Supplementray Tables

[Table 1 about here.]

[Table 2 about here.]

[Table 3 about here.]

References

Jackel, P. (2005). A note on multivariate Gauss-Hermite quadrature. *London: ABN-Amro. Re.*

Received XX 20XX. *Revised XX* 20XX. *Accepted XX* 20XX.

Table S1

*Comparison of operating characteristics for the naive method ($\hat{\beta}^{(N)}$), approximate calibration method ($\hat{\beta}^{(A)}$), and Monte Carlo and GHQ exact calibration methods ($\hat{\beta}^{(E1)}$ and $\hat{\beta}^{(E2)}$), assuming $\boldsymbol{Z} = W$ is included in the model for disease outcome (**??**) with regression coefficient $\beta_z = log(1.25)$.*

β_x	Percent Bias				MSE($\times$100)				SE($\times$100)				Coverage Rate($\times$100)			
	$\hat{\beta}_x^{(N)}$	$\hat{\beta}_x^{(A)}$	$\hat{\beta}_x^{(E1)}$	$\hat{\beta}_x^{(E2)}$	$\hat{\beta}_x^{(N)}$	$\hat{\beta}_x^{(A)}$	$\hat{\beta}_x^{(E1)}$	$\hat{\beta}_x^{(E2)}$	$\hat{\beta}_x^{(N)}$	$\hat{\beta}_x^{(A)}$	$\hat{\beta}_x^{(E1)}$	$\hat{\beta}_x^{(E2)}$	$\hat{\beta}_x^{(N)}$	$\hat{\beta}_x^{(A)}$	$\hat{\beta}_x^{(E1)}$	$\hat{\beta}_x^{(E2)}$
log(1.25)	-32.5	1.9	0.5	2.4	0.98	0.76	0.75	0.78	6.76	8.70	8.68	8.81	57.5	94.6	94.3	94.6
log(1.5)	-34.0	-0.0	-1.0	1.2	2.48	1.10	1.13	1.19	7.62	10.52	10.62	10.91	30.5	92.6	92.2	91.9
log(1.75)	-33.1	0.6	0.4	2.6	4.00	1.29	1.38	1.49	7.64	11.34	11.76	12.12	14.2	90.2	89.5	89.5
log(2)	-34.1	0.1	0.5	3.0	6.22	1.56	1.71	1.96	7.92	12.51	13.09	13.84	6.7	87.8	87.2	86.9

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.

Table S2

Comparison of operating characteristics for the naive method ($\hat{\beta}^{(N)}$), full calibration method ($\hat{\beta}^{(F)}$), approximate calibration method ($\hat{\beta}^{(A)}$), and Monte Carlo and GHQ exact calibration methods ($\hat{\beta}^{(E1)}$ and $\hat{\beta}^{(E2)}$), for the standard deviation of γ_d ranging from 0.05 to 0.15.

σ_γ	β_x	Percent Bias					MSE(×100)					SE(×100)					Coverage Rate(×100)				
		$\hat{\beta}_x^{(N)}$	$\hat{\beta}_x^{(F)}$	$\hat{\beta}_x^{(A)}$	$\hat{\beta}_x^{(E1)}$	$\hat{\beta}_x^{(E2)}$	$\hat{\beta}_x^{(N)}$	$\hat{\beta}_x^{(F)}$	$\hat{\beta}_x^{(A)}$	$\hat{\beta}_x^{(E1)}$	$\hat{\beta}_x^{(E2)}$	$\hat{\beta}_x^{(N)}$	$\hat{\beta}_x^{(F)}$	$\hat{\beta}_x^{(A)}$	$\hat{\beta}_x^{(E1)}$	$\hat{\beta}_x^{(E2)}$	$\hat{\beta}_x^{(N)}$	$\hat{\beta}_x^{(F)}$	$\hat{\beta}_x^{(A)}$	$\hat{\beta}_x^{(E1)}$	$\hat{\beta}_x^{(E2)}$
0.05	log(1.25)	-19.0	-1.1	-0.3	-0.5	0.1	0.38	0.31	0.28	0.29	0.29	4.51	5.54	5.31	5.36	5.38	79.4	93.3	95.6	95.2	95.4
	log(1.5)	-19.8	-1.6	-1.2	-0.8	0.0	0.86	0.40	0.31	0.33	0.33	4.66	6.27	5.58	5.75	5.79	54.6	90.9	95.8	95.2	95.6
	log(1.75)	-20.5	-1.7	-1.6	-0.4	0.5	1.55	0.48	0.38	0.41	0.42	4.88	6.87	6.08	6.41	6.50	32.5	90.0	94.5	94.3	94.5
	log(2)	-21.1	-1.5	-2.3	-0.2	0.8	2.41	0.69	0.44	0.49	0.51	5.18	8.24	6.47	7.01	7.15	17.8	86.5	93.5	93.0	93.2
0.10	log(1.25)	-19.3	-0.5	1.0	0.7	1.4	0.43	0.34	0.28	0.28	0.28	4.93	5.83	5.26	5.29	5.32	77.2	91.9	96.4	96.5	96.1
	log(1.5)	-19.8	-0.4	0.4	0.7	1.6	0.92	0.54	0.33	0.35	0.36	5.20	7.34	5.75	5.93	5.97	52.0	86.2	95.3	94.9	95.0
	log(1.75)	-20.8	-0.5	-0.7	0.4	1.4	1.66	0.79	0.41	0.45	0.47	5.55	8.87	6.42	6.71	6.85	32.3	80.7	93.6	93.9	93.2
	log(2)	-21.6	-0.9	-1.7	0.4	1.4	2.58	1.05	0.48	0.56	0.58	5.83	10.24	6.83	7.46	7.55	18.3	78.9	91.8	92.1	92.1
0.15	log(1.25)	-21.1	0.2	2.7	2.4	3.2	0.53	0.44	0.30	0.31	0.31	5.55	6.61	5.47	5.51	5.53	71.1	88.8	95.5	95.2	95.0
	log(1.5)	-21.4	0.9	1.3	1.6	2.5	1.11	0.88	0.40	0.41	0.43	5.98	9.40	6.28	6.41	6.50	46.8	78.1	95.2	94.8	94.4
	log(1.75)	-22.1	0.4	0.8	2.0	2.9	1.95	1.41	0.53	0.60	0.63	6.56	11.87	7.30	7.68	7.77	29.4	70.6	92.0	90.9	90.0
	log(2)	-22.1	0.9	0.4	2.4	3.5	2.82	1.81	0.59	0.71	0.77	6.90	13.46	7.69	8.25	8.42	21.4	68.2	90.6	89.5	88.8

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.

Table S3

Comparison of operating characteristics for the naive method ($\hat{\beta}^{(N)}$), full calibration method ($\hat{\beta}^{(F)}$), approximate calibration method ($\hat{\beta}^{(A)}$), and Monte Carlo and GHQ exact calibration methods ($\hat{\beta}^{(E1)}$ and $\hat{\beta}^{(E2)}$), where 30% of controls in each contribution study were selected in the calibration subset.

β_x	Percent Bias					MSE(×100)					SE(×100)					Coverage Rate(×100)				
	$\hat{\beta}_x^{(N)}$	$\hat{\beta}_x^{(F)}$	$\hat{\beta}_x^{(A)}$	$\hat{\beta}_x^{(E1)}$	$\hat{\beta}_x^{(E2)}$	$\hat{\beta}_x^{(N)}$	$\hat{\beta}_x^{(F)}$	$\hat{\beta}_x^{(A)}$	$\hat{\beta}_x^{(E1)}$	$\hat{\beta}_x^{(E2)}$	$\hat{\beta}_x^{(N)}$	$\hat{\beta}_x^{(F)}$	$\hat{\beta}_x^{(A)}$	$\hat{\beta}_x^{(E1)}$	$\hat{\beta}_x^{(E2)}$	$\hat{\beta}_x^{(N)}$	$\hat{\beta}_x^{(F)}$	$\hat{\beta}_x^{(A)}$	$\hat{\beta}_x^{(E1)}$	$\hat{\beta}_x^{(E2)}$
log(1.25)	-18.3	1.8	-0.3	-0.6	0.0	1.04	0.34	0.26	0.26	0.26	9.34	5.83	5.09	5.11	5.14	56.5	92.1	97.6	97.6	97.5
log(1.5)	-18.1	1.2	-2.6	-2.3	-1.6	1.47	0.49	0.32	0.33	0.33	9.62	7.00	5.55	5.69	5.72	49.8	89.4	97.3	97.0	97.2
log(1.75)	-19.8	0.3	-3.4	-2.4	-1.6	2.15	0.67	0.42	0.44	0.43	9.57	8.18	6.16	6.49	6.51	39.6	85.0	95.5	95.3	95.5
log(2)	-20.0	0.6	-4.0	-2.2	-1.4	3.10	1.04	0.55	0.57	0.56	10.83	10.18	6.87	7.38	7.45	33.9	78.3	93.6	93.3	93.9

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.