Supplementary Material for "Biomarker Combinations for Risk Prediction in Multicenter Studies: Principles and Methods"

Allison Meisner[1], Chirag R. Parikh[2,3], and Kathleen F. Kerr[1]

[1]Department of Biostatistics, University of Washington, Seattle, Washington

[2]Program of Applied Translational Research, Department of Medicine, Yale School of Medicine, New Haven, Connecticut

[3]Department of Internal Medicine, Veterans Affairs Medical Center, West Haven, Connecticut

meisnera@uw.edu

S1 Risk Function for Conditionally Bivariate Normal Biomarkers

Claim: If the biomarkers X_1 and X_2 have the conditional distribution given by equation (6), the true risk function is given by

$$\text{logit}\{P(D=1|X_1, X_2, C=c)\} = \beta_0^c + \beta_1 X_1 + \beta_2 X_2,$$

where

$$\beta_0^c = \frac{-\mu_{X_1}^2 - \mu_{X_2}^2}{2(1-\rho^2)} + \frac{\rho\mu_{X_1}\mu_{X_2} + \rho\mu_{X_1} f_{X_2}(c) + \rho\mu_{X_2} f_{X_1}(c) - \mu_{X_1} f_{X_1}(c) - \mu_{X_2} f_{X_2}(c)}{1-\rho^2}$$
$$+ \log\left(\frac{\gamma_c}{1-\gamma_c}\right)$$
$$\beta_1 = \frac{\mu_{X_1} - \rho\mu_{X_2}}{1-\rho^2}$$
$$\beta_2 = \frac{\mu_{X_2} - \rho\mu_{X_1}}{1-\rho^2}.$$

Proof. We can demonstrate this as follows:

$$\begin{aligned} P(D=1|X_1, X_2, C=c) &= \frac{f(X_1, X_2|D=1, C=c)\gamma_c}{f(X_1, X_2|D=1, C=c)\gamma_c + f(X_1, X_2|D=0, C=c)(1-\gamma_c)} \\ &= \frac{1}{1+B/A}, \end{aligned}$$

where $A = f(X_1, X_2|D=1, C=c)\gamma_c$ and $B = f(X_1, X_2|D=0, C=c)(1-\gamma_c)$. We have

$$\begin{aligned} \frac{B}{A} &= \frac{f(X_1, X_2|D=0, C=c)(1-\gamma_c)}{f(X_1, X_2|D=1, C=c)\gamma_c} \\ &= \exp\Bigg(\frac{1}{2(1-\rho^2)}\big[\{X_1 - \mu_{X_1} - f_{X_1}(c)\}^2 - \{X_1 - f_{X_1}(c)\}^2 \quad \times \left(\frac{1-\gamma_c}{\gamma_c}\right) \\ &\qquad - 2\rho\{X_1 - \mu_{X_1} - f_{X_1}(c)\}\{X_2 - \mu_{X_2} - f_{X_2}(c)\} \\ &\qquad\qquad + 2\rho\{X_1 - f_{X_1}(c)\}\{X_2 - f_{X_2}(c)\} \\ &\qquad\qquad + \{X_2 - \mu_{X_2} - f_{X_2}(c)\}^2 - \{X_2 - f_{X_2}(c)\}^2\big]\Bigg) \\ &= \exp\Bigg\{\frac{1}{2(1-\rho^2)}\big(\mu_{X_1}^2 - 2\mu_{X_1}\{X_1 - f_{X_1}(c)\} + 2\rho[-\mu_{X_1}\mu_{X_2} + \mu_{X_1}\{X_2 - f_{X_2}(c)\} + \\ &\qquad \mu_{X_2}\{X_1 - f_{X_1}(c)\}] + \mu_{X_2}^2 - 2\mu_{X_2}\{X_2 - f_{X_2}(c)\}\big) + \log\left(\frac{1-\gamma_c}{\gamma_c}\right)\Bigg\}. \end{aligned}$$

If $P(D=1|X_1,\,X_2,\,C=c)=\dfrac{1}{1+\exp(\star)}$ then $P(D=1|X_1,X_2,C=c)=\text{expit}(-\star)$, so

$$
\begin{aligned}
&P(D=1|X_1,X_2,C=c)\\
&\quad=\text{expit}\Bigg\{\frac{-1}{2(1-\rho^2)}\big(\mu_{X_1}^2-2\mu_{X_1}\{X_1-f_{X_1}(c)\}\\
&\qquad+2\rho[-\mu_{X_1}\mu_{X_2}+\mu_{X_1}\{X_2-f_{X_2}(c)\}+\mu_{X_2}\{X_1-f_{X_1}(c)\}]\\
&\qquad\quad+\mu_{X_2}^2-2\mu_{X_2}\{X_2-f_{X_2}(c)\}\big)-\log\left(\frac{1-\gamma_c}{\gamma_c}\right)\Bigg\}\\
&\quad=\text{expit}\Bigg\{\frac{-\mu_{X_1}^2-\mu_{X_2}^2}{2(1-\rho^2)}+\log\left(\frac{\gamma_c}{1-\gamma_c}\right)+\frac{\mu_{X_1}-\rho\mu_{X_2}}{1-\rho^2}X_1+\frac{\mu_{X_2}-\rho\mu_{X_1}}{1-\rho^2}X_2\\
&\qquad+\frac{\rho\mu_{X_1}\mu_{X_2}+\rho\mu_{X_1}f_{X_2}(c)+\rho\mu_{X_2}f_{X_1}(c)-\mu_{X_1}f_{X_1}(c)-\mu_{X_2}f_{X_2}(c)}{1-\rho^2}\Bigg\}\\
&\quad=\text{expit}(\beta_0^c+\beta_1X_1+\beta_2X_2).
\end{aligned}
$$

□

S2 Lemma 1, Theorem 1, and Theorem 2

First, we describe the conditions required for these results to hold.

(C1) The m centers are randomly sampled from the population of M centers, and n_c observations are randomly sampled from center c, $c = 1, ..., m$.

(C2) $\sum_{c=1}^{m} |E(\hat{w}_c) - w_c| \to 0$ as $n_c \to \infty$, $c = 1, ..., m$, and $m \to M$ such that $\sqrt{n_c}/m \to \infty$.

(C3) The centers are independent and within each center, the observations $O_i^c = (D_i^c, \mathbf{X}_i^c)$, $i = 1, ..., n_c$, are independent and identically distributed $(p + 1)$-dimensional random vectors such that there exists at least one component of $\mathbf{X}^c$, X_k^c for some $k \in \{1, ..., p\}$, with distribution that has everywhere positive Lebesgue density, conditional on the other $\mathbf{X}^c$ components.

(C4) The support of $\mathbf{X}^c$, $c = 1, ..., M$, is not contained in any proper linear subspace of $\mathbb{R}^p$.

(C5) $AUC_c(\boldsymbol{\theta})$ is differentiable at $\boldsymbol{\theta}_0$ and $||AUC_c'(\boldsymbol{\theta}_0)|| \leq T < \infty$, $c = 1, ..., m$.

(C6) $\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}_0 \in B = \{\boldsymbol{\theta} \in \mathbb{R}^p : ||\boldsymbol{\theta}|| = 1, |\theta_k| > 0\}$

In Lemma 1, Theorem 1, and Theorem 2, we restrict the combinations to have $||\boldsymbol{\theta}|| = 1$ for mathematical ease; since AUC_c is invariant to monotone increasing transformations, this is not restrictive in a practical sense. We will assume a non-trivial disease prevalence throughout; in particular, we assume $P(D = 1|C = c) := \gamma_c \in [1/V, 1 - 1/V]$, $c = 1, ..., M$, for some $V \in (2, \infty)$.

Lemma 1. *Suppose conditions (C1), (C3), and (C4) hold for a given center c. Then*

$\sup_{\boldsymbol{\theta} \in B} \left| A\hat{U}C_c(\boldsymbol{\theta}) - AUC_c(\boldsymbol{\theta}) \right| = o_p(1)$ *as* $n_c \to \infty$*, where* $B = \{\boldsymbol{\theta} \in \mathbb{R}^p : ||\boldsymbol{\theta}|| = 1, |\theta_k| > 0\}$.

Proof. Previous work by Han[1] proved a similar claim for a related statistic. Consider a single center with n total observations, n_D cases, and $n_{\bar{D}}$ controls. Let $\mathbf{X}_i$ denote the biomarker vector for observation i. Han considered the statistic $h_{ij}(\boldsymbol{\theta})$: for any pair of observations (i, j), let

$$h_{ij}(\boldsymbol{\theta}) = 1(D_i > D_j)1(\boldsymbol{\theta}^\top \mathbf{X}_i > \boldsymbol{\theta}^\top \mathbf{X}_j) + 1(D_i < D_j)1(\boldsymbol{\theta}^\top \mathbf{X}_i < \boldsymbol{\theta}^\top \mathbf{X}_j).$$

Then

$$\begin{aligned}
\sum_{\kappa} h_{ij}(\boldsymbol{\theta}) &= \sum_{i<j} h_{ij}(\boldsymbol{\theta}) \\
&= \frac{1}{2} \sum_{i \neq j} 1\{(D_i - D_j)(\boldsymbol{\theta}^\top \mathbf{X}_i - \boldsymbol{\theta}^\top \mathbf{X}_j) > 0\} \\
&= \sum_{i \neq j} 1(D_i > D_j)1(\boldsymbol{\theta}^\top \mathbf{X}_i > \boldsymbol{\theta}^\top \mathbf{X}_j) \\
&= \sum_{i=1}^{n_D} \sum_{j=1}^{n_{\bar{D}}} 1(\boldsymbol{\theta}^\top \mathbf{X}_{Di} > \boldsymbol{\theta}^\top \mathbf{X}_{\bar{D}j}),
\end{aligned}$$

where κ denotes the collection of all pairs of distinct elements (regardless of D). Also, for any $i \neq j$,

$$\begin{aligned}
P(\boldsymbol{\theta}^\top \mathbf{X}_i < \boldsymbol{\theta}^\top \mathbf{X}_j, D_i < D_j) + P(\boldsymbol{\theta}^\top \mathbf{X}_i > \boldsymbol{\theta}^\top \mathbf{X}_j, D_i > D_j) &= P(\boldsymbol{\theta}^\top \mathbf{X}_i < \boldsymbol{\theta}^\top \mathbf{X}_j | D_i < D_j) P(D_i < D_j) \\
&\quad + P(\boldsymbol{\theta}^\top \mathbf{X}_i > \boldsymbol{\theta}^\top \mathbf{X}_j | D_i > D_j) P(D_i > D_j) \\
&= P(\boldsymbol{\theta}^\top \mathbf{X}_i < \boldsymbol{\theta}^\top \mathbf{X}_j | D_i < D_j) \gamma(1-\gamma) \\
&\quad + P(\boldsymbol{\theta}^\top \mathbf{X}_i > \boldsymbol{\theta}^\top \mathbf{X}_j | D_i > D_j) \gamma(1-\gamma) \\
&= 2\gamma(1-\gamma) AUC(\boldsymbol{\theta}),
\end{aligned}$$

where γ is the prevalence.

We can consider

$$\begin{aligned}
S_n(\boldsymbol{\theta}) &= \frac{2}{n(n-1)} \sum_{\kappa} h_{ij}(\boldsymbol{\theta}) \\
A\hat{U}C(\boldsymbol{\theta}) &= \frac{1}{n_D n_{\bar{D}}} \sum_{i=1}^{n_D} \sum_{j=1}^{n_{\bar{D}}} 1(\boldsymbol{\theta}^\top \mathbf{X}_{Di} > \boldsymbol{\theta}^\top \mathbf{X}_{\bar{D}j}),
\end{aligned}$$

where S_n is a one-sample U-statistic and $A\hat{U}C$ is a two-sample U-statistic. We would like to be able to study the asymptotic behavior of S_n; in particular, we would like to say that if

$$\sup_{\boldsymbol{\theta} \in B} |S_n(\boldsymbol{\theta}) - 2\gamma(1-\gamma) AUC(\boldsymbol{\theta})| = o_p(1)$$

then $\sup_{\boldsymbol{\theta} \in B} \left| A\hat{U}C(\boldsymbol{\theta}) - AUC(\boldsymbol{\theta}) \right| = o_p(1)$.

If $\sup_{\boldsymbol{\theta}\in B} |S_n(\boldsymbol{\theta}) - 2\gamma(1-\gamma)AUC(\boldsymbol{\theta})| = o_p(1)$, then

$$\sup_{\boldsymbol{\theta}\in B}\left|\frac{S_n(\boldsymbol{\theta})}{2\gamma(1-\gamma)} - AUC(\boldsymbol{\theta})\right| = o_p(1).$$

We would then have

$$\begin{aligned}
\sup_{\boldsymbol{\theta}\in B}\left|A\hat{U}C(\boldsymbol{\theta}) - AUC(\boldsymbol{\theta})\right| &\leq \sup_{\boldsymbol{\theta}\in B}\left|A\hat{U}C(\boldsymbol{\theta}) - \frac{S_n(\boldsymbol{\theta})}{2\gamma(1-\gamma)}\right| + \sup_{\boldsymbol{\theta}\in B}\left|\frac{S_n(\boldsymbol{\theta})}{2\gamma(1-\gamma)} - AUC(\boldsymbol{\theta})\right| \\
&= \sup_{\boldsymbol{\theta}\in B}\left|A\hat{U}C(\boldsymbol{\theta}) - \frac{S_n(\boldsymbol{\theta})}{2\gamma(1-\gamma)}\right| + o_p(1) \\
&\leq \sup_{\boldsymbol{\theta}\in B}\left|A\hat{U}C(\boldsymbol{\theta}) - \frac{S_n(\boldsymbol{\theta})n^2}{2n_D n_{\bar{D}}}\right| + \sup_{\boldsymbol{\theta}\in B}\left|\frac{S_n(\boldsymbol{\theta})n^2}{2n_D n_{\bar{D}}} - \frac{S_n(\boldsymbol{\theta})}{2\gamma(1-\gamma)}\right| + o_p(1) \\
&= \left|1 - \frac{n^2}{n(n-1)}\right| \sup_{\boldsymbol{\theta}\in B} A\hat{U}C(\boldsymbol{\theta}) + \sup_{\boldsymbol{\theta}\in B}\left|\frac{S_n(\boldsymbol{\theta})n^2}{2n_D n_{\bar{D}}} - \frac{S_n(\boldsymbol{\theta})}{2\gamma(1-\gamma)}\right| + o_p(1) \\
&\leq o(1) + \left|\frac{n^2}{2n_D n_{\bar{D}}} - \frac{1}{2\gamma(1-\gamma)}\right| + o_p(1) \\
&= o_p(1),
\end{aligned}$$

where the last inequality follows from the fact that $S_n(\boldsymbol{\theta}), A\hat{U}C(\boldsymbol{\theta}) \leq 1$ and the last equality follows from the Weak Law of Large Numbers, the continuous mapping theorem and Slutsky's theorem. Thus, to demonstrate uniform convergence of $A\hat{U}C(\boldsymbol{\theta})$, we can consider $S_n(\boldsymbol{\theta})$.

The following is taken nearly verbatim (with very minor variations) from part of the proof given in Han.[1] Let

$$
\begin{aligned}
h(\boldsymbol{\theta}) &= E\{h_{ij}(\boldsymbol{\theta})\} \equiv 2\gamma(1-\gamma)AUC(\boldsymbol{\theta}) \\
\bar{g}_{ij}(\boldsymbol{\theta},\delta) &= \sup_{b \in D_\delta(\boldsymbol{\theta})} \{h_{ij}(b) - h(b)\} \\
\underline{g}_{ij}(\boldsymbol{\theta},\delta) &= \inf_{b \in D_\delta(\boldsymbol{\theta})} \{h_{ij}(b) - h(b)\} \\
\bar{g}(\boldsymbol{\theta},\delta) &= E\{\bar{g}_{ij}(\boldsymbol{\theta},\delta)\} \\
\underline{g}(\boldsymbol{\theta},\delta) &= E\{\underline{g}_{ij}(\boldsymbol{\theta},\delta)\}
\end{aligned}
$$

where $D_\delta(\boldsymbol{\theta}) = \{b : b \in B, ||b - \boldsymbol{\theta}|| < \delta\}$.

It can be seen that $h_{ij}(\boldsymbol{\theta})$ is a step function uniformly bounded in (i,j) and $\boldsymbol{\theta}$. By condition (C3), $h_{ij}(\boldsymbol{\theta})$ is continuous in $\boldsymbol{\theta} \in B$ uniformly across (i,j) almost surely. Thus, $h(\boldsymbol{\theta})$ is uniformly bounded and continuous in $\boldsymbol{\theta} \in B$.

Note also that $\bar{g}_{ij}(\boldsymbol{\theta},\delta)$ is measurable for all $\boldsymbol{\theta} \in B$ and $\delta > 0$ since B is separable and for any $\boldsymbol{\theta} \in B$ there exists a sequence $\{\boldsymbol{\theta}_t\}$ in a countable dense subset of B such that

$$
\lim_{t\to\infty} h_{ij}(\boldsymbol{\theta}_t) = h_{ij}(\boldsymbol{\theta}), \quad \lim_{t\to\infty} h(\boldsymbol{\theta}_t) = h(\boldsymbol{\theta}).
$$

Also, $\bar{g}_{ij}(\boldsymbol{\theta},\delta)$ is uniformly bounded in $\boldsymbol{\theta}$ and

$$
\lim_{\delta\to 0} \bar{g}_{ij}(\boldsymbol{\theta},\delta) = h_{ij}(\boldsymbol{\theta}) - h(\boldsymbol{\theta}) \text{ almost surely.}
$$

Thus, it follows that $\lim_{\delta\to 0} \bar{g}(\boldsymbol{\theta}, \delta) = 0$ for all $\boldsymbol{\theta} \in B$. A similar argument can be made for $\underline{g}(\boldsymbol{\theta}, \delta)$, giving $\lim_{\delta\to 0} \underline{g}(\boldsymbol{\theta}, \delta) = 0$ for all $\boldsymbol{\theta} \in B$.

To show convergence of $S_n(\boldsymbol{\theta})$ to $h(\boldsymbol{\theta})$ uniformly in $\boldsymbol{\theta}$, we have for a given $\epsilon > 0$,

$$
\begin{aligned}
P\left(\sup_{\boldsymbol{\theta}\in B} |S_n(\boldsymbol{\theta}) - h(\boldsymbol{\theta})| > \epsilon\right) \leq & P\left(\left|\frac{2}{n(n-1)} \sum_{\kappa} \sup_{\boldsymbol{\theta}\in B} \{h_{ij}(\boldsymbol{\theta}) - h(\boldsymbol{\theta})\}\right| > \epsilon\right) \\
& + P\left(\left|\frac{2}{n(n-1)} \sum_{\kappa} \inf_{\boldsymbol{\theta}\in B} \{h_{ij}(\boldsymbol{\theta}) - h(\boldsymbol{\theta})\}\right| > \epsilon\right).
\end{aligned}
$$

Since B is compact, there exists a finite set of coverings $\{D_{\delta_l}(\boldsymbol{\theta}_l)\}$, $l = 1, ..., L$, such that

$$
B \subset \bigcup_{l=1}^{L} D_{\delta_l}(\boldsymbol{\theta}_l) \text{ and } \bar{g}(\boldsymbol{\theta}_l, \delta_l), \underline{g}(\boldsymbol{\theta}_l, \delta_l) > \epsilon/2.
$$

Thus,

$$
\begin{aligned}
P\left(\sup_{\boldsymbol{\theta}\in B} |S_n(\boldsymbol{\theta}) - h(\boldsymbol{\theta})| > \epsilon\right) \leq & \sum_{l=1}^{L} P\left(\left|\frac{2}{n(n-1)} \sum_{\kappa} \bar{g}_{ij}(\boldsymbol{\theta}_l, \delta_l) - \bar{g}(\boldsymbol{\theta}_l, \delta_l)\right| \geq \epsilon/2\right) \\
& + \sum_{l=1}^{L} P\left(\left|\frac{2}{n(n-1)} \sum_{\kappa} \underline{g}_{ij}(\boldsymbol{\theta}_l, \delta_l) - \underline{g}(\boldsymbol{\theta}_l, \delta_l)\right| \geq \epsilon/2\right).
\end{aligned}
$$

However, for each l, $\frac{2}{n(n-1)} \sum_{\kappa} \bar{g}_{ij}(\boldsymbol{\theta}_l, \delta_l)$ is a one-sample U-statistic with kernel

$$
\bar{g}_{ij}(\boldsymbol{\theta}_l, \delta_l) = \sup_{\mathbf{b}\in D_{\delta_l}(\boldsymbol{\theta}_l)} \{h_{ij}(\mathbf{b}) - h(\mathbf{b})\}.
$$

Since $E\{\bar{g}_{ij}(\boldsymbol{\theta}_l, \delta_l)\}^2 \leq E(1) < \infty$, we have

$$
\frac{2}{n(n-1)} \sum_{\kappa} \bar{g}_{ij}(\boldsymbol{\theta}_l, \delta_l) \xrightarrow{p} \bar{g}(\boldsymbol{\theta}_l, \delta_l).
$$

These arguments also apply to $\underline{g}_{ij}(\boldsymbol{\theta}_l, \delta_l)$, giving

$$\frac{2}{n(n-1)} \sum_{\kappa} \bar{g}_{ij}(\boldsymbol{\theta}_l, \delta_l) \xrightarrow{p} \bar{g}(\boldsymbol{\theta}_l, \delta_l), \quad l = 1, ..., L$$
$$\frac{2}{n(n-1)} \sum_{\kappa} \underline{g}_{ij}(\boldsymbol{\theta}_l, \delta_l) \xrightarrow{p} \underline{g}(\boldsymbol{\theta}_l, \delta_l), \quad l = 1, ..., L.$$

Thus, we have

$$P\left(\sup_{\boldsymbol{\theta} \in B} |S_n(\boldsymbol{\theta}) - h(\boldsymbol{\theta})| > \epsilon\right) \to 0,$$

so $\sup_{\boldsymbol{\theta} \in B} |A\hat{U}C(\boldsymbol{\theta}) - AUC(\boldsymbol{\theta})| \xrightarrow{p} 0$. Since this holds for any center c, we have demonstrated $\sup_{\boldsymbol{\theta} \in B} |A\hat{U}C_c(\boldsymbol{\theta}) - AUC_c(\boldsymbol{\theta})| \xrightarrow{p} 0$ as $n_c \to \infty$. □

Theorem 1. *Suppose conditions (C1)–(C4) hold. Then for* $B = \{\boldsymbol{\theta} \in \mathbb{R}^p : ||\boldsymbol{\theta}|| = 1, |\theta_k| > 0\}$,

$$\sup_{\boldsymbol{\theta} \in B} \left| a\hat{AUC}(\boldsymbol{\theta}) - aAUC(\boldsymbol{\theta}) \right| \xrightarrow{p} 0$$

and

$$\sum_{c=1}^{m} |\hat{w}_c - w_c| = o_p(1)$$

as $n_c \to \infty$, $c = 1, ..., m$, *and* $m \to M$ *such that* $\sqrt{n_c}/m \to \infty$.

Proof. We can write this claim as follows:

$$\sup_{\boldsymbol{\theta} \in B} |a\hat{AUC}(\boldsymbol{\theta}) - aAUC(\boldsymbol{\theta})| = \sup_{\boldsymbol{\theta} \in B} \left| \sum_{c=1}^{m} \hat{w}_c \hat{AUC}_c(\boldsymbol{\theta}) - \sum_{c=1}^{M} w_c AUC_c(\boldsymbol{\theta}) \right| \xrightarrow{p} 0.$$

We can write

$$\begin{aligned}
&\sup_{\boldsymbol{\theta} \in B} \left| \sum_{c=1}^{m} \hat{w}_c \hat{AUC}_c(\boldsymbol{\theta}) - \sum_{c=1}^{M} w_c AUC_c(\boldsymbol{\theta}) \right| \\
&\quad = \sup_{\boldsymbol{\theta} \in B} \left| \sum_{c=1}^{m} \hat{w}_c \hat{AUC}_c(\boldsymbol{\theta}) - \sum_{c=1}^{m} w_c AUC_c(\boldsymbol{\theta}) - \sum_{c=m+1}^{M} w_c AUC_c(\boldsymbol{\theta}) \right| \\
&\quad = \sup_{\boldsymbol{\theta} \in B} \left| \sum_{c=1}^{m} \left(\hat{w}_c \hat{AUC}_c(\boldsymbol{\theta}) - w_c AUC_c(\boldsymbol{\theta}) \right) - \sum_{c=m+1}^{M} w_c AUC_c(\boldsymbol{\theta}) \right| \\
&\quad \leq \sup_{\boldsymbol{\theta} \in B} \left| \sum_{c=1}^{m} \left(\hat{w}_c \hat{AUC}_c(\boldsymbol{\theta}) - w_c AUC_c(\boldsymbol{\theta}) \right) \right| + \sum_{c=m+1}^{M} \sup_{\boldsymbol{\theta} \in B} |w_c AUC_c(\boldsymbol{\theta})| \\
&\quad = \sup_{\boldsymbol{\theta} \in B} \left| \sum_{c=1}^{m} \left(\hat{w}_c \hat{AUC}_c(\boldsymbol{\theta}) - w_c AUC_c(\boldsymbol{\theta}) \right) \right| + o(1)
\end{aligned}$$

as $m \to M$.

Then

$$
\begin{aligned}
&\sup_{\boldsymbol{\theta}\in B}\left|\sum_{c=1}^{m}\left(\hat{w}_c A\hat{U}C_c(\boldsymbol{\theta}) - w_c AUC_c(\boldsymbol{\theta})\right)\right| \\
&\quad= \sup_{\boldsymbol{\theta}\in B}\left|\sum_{c=1}^{m}\left\{\hat{w}_c\left(A\hat{U}C_c(\boldsymbol{\theta}) - AUC_c(\boldsymbol{\theta})\right) + \hat{w}_c AUC_c(\boldsymbol{\theta}) - w_c AUC_c(\boldsymbol{\theta})\right\}\right| \\
&\quad\le \sup_{\boldsymbol{\theta}\in B}\left|\sum_{c=1}^{m}\hat{w}_c\left(A\hat{U}C_c(\boldsymbol{\theta}) - AUC_c(\boldsymbol{\theta})\right)\right| + \sup_{\boldsymbol{\theta}\in B}\left|\sum_{c=1}^{m}\left(\hat{w}_c - w_c\right)AUC_c(\boldsymbol{\theta})\right| \\
&\quad\le \sum_{c=1}^{m}\sup_{\boldsymbol{\theta}\in B}\left|\hat{w}_c\left(A\hat{U}C_c(\boldsymbol{\theta}) - AUC_c(\boldsymbol{\theta})\right)\right| + \sum_{c=1}^{m}\sup_{\boldsymbol{\theta}\in B}\left|\left(\hat{w}_c - w_c\right)AUC_c(\boldsymbol{\theta})\right| \\
&\quad= \sum_{c=1}^{m}\hat{w}_c\sup_{\boldsymbol{\theta}\in B}\left|A\hat{U}C_c(\boldsymbol{\theta}) - AUC_c(\boldsymbol{\theta})\right| + \sum_{c=1}^{m}\left|\hat{w}_c - w_c\right|\sup_{\boldsymbol{\theta}\in B}AUC_c(\boldsymbol{\theta}) \\
&\quad= \sum_{c=1}^{m}\hat{w}_c o_p(1) + \sum_{c=1}^{m}\left|\hat{w}_c - w_c\right|\sup_{\boldsymbol{\theta}\in B}AUC_c(\boldsymbol{\theta}) \\
&\quad= o_p(1) + \sum_{c=1}^{m}\left|\hat{w}_c - w_c\right|\sup_{\boldsymbol{\theta}\in B}AUC_c(\boldsymbol{\theta}) \\
&\quad\le o_p(1) + \sum_{c=1}^{m}\left|\hat{w}_c - w_c\right|,
\end{aligned}
$$

where the second to last equality follows from the fact that $\sup_{\boldsymbol{\theta}\in B}\left|A\hat{U}C_c(\boldsymbol{\theta}) - AUC_c(\boldsymbol{\theta})\right| = o_p(1)$ by Lemma 1, the last equality follows from the fact that $\sum_{c=1}^{m}\hat{w}_c = 1$ for every m, and the last inequality follows from $AUC_c(\boldsymbol{\theta}) \le 1$.

Now we must show that

$$\sum_{c=1}^{m}\left|\hat{w}_c - w_c\right| = o_p(1).$$

Equivalently, we must prove that for every $\epsilon > 0$,

$$P\left(\sum_{c=1}^{m}\left|\hat{w}_c - w_c\right| > \epsilon\right) \to 0$$

as $n_c \to \infty$, $c = 1, ..., m$, and $m \to M$ such that $\sqrt{n_c}/m \to \infty$.

We have

$$P\left(\sum_{c=1}^{m} |\hat{w}_c - w_c| > \epsilon\right) \leq \frac{E(\sum_{c=1}^{m} |\hat{w}_c - w_c|)}{\epsilon} = \frac{\sum_{c=1}^{m} E|\hat{w}_c - w_c|}{\epsilon}$$
$$\leq \frac{\sum_{c=1}^{m} E\,|\hat{w}_c - E(\hat{w}_c)|}{\epsilon} + \frac{\sum_{c=1}^{m} |E(\hat{w}_c) - w_c|}{\epsilon}$$
$$\leq \frac{\sum_{c=1}^{m} \sqrt{Var(\hat{w}_c)}}{\epsilon} + o(1),$$

where the first inequality follows from Markov's inequality and the third inequality follows from Jensen's inequality and the fact that $\sum_{c=1}^{m} |E(\hat{w}_c) - w_c| = o(1)$ by condition (C2).

Let $f(R, S) = R/S$ and $\nu = (E(R), E(S))$. Then we can use a first-order Taylor approximation to write

$$f(R, S) \approx f(\nu) + f'_R(\nu)(R - E(R)) + f'_S(\nu)(S - E(S)),$$

where $f'_R(\nu) = \frac{\partial f(R,S)}{\partial R}|_\nu$ and $f'_S(\nu)$ is defined analogously. This gives $E[f(R, S)] \approx f(\nu)$. We can also write

$$Var[f(R, S)] \approx E\left[\{f(R, S) - f(\nu)\}^2\right] \approx \left[\frac{1}{\{E(S)\}^2}\right] Var(R) + \left[\frac{\{E(R)\}^2}{\{E(S)\}^4}\right] Var(S)$$
$$- 2\left[\frac{E(R)}{\{E(S)\}^3}\right] Cov(R, S).$$

In our case, we have $\hat{w}_c = n_D^c/n_D$, giving $R = n_D^c, S = n_D$, so $E(R) = n_c\gamma_c$, $Var(R) = n_c\gamma_c(1-\gamma_c)$, $E(S) = \sum_{c=1}^m n_c\gamma_c$, $Var(S) = \sum_{c=1}^m n_c\gamma_c(1-\gamma_c)$, and $Cov(R,S) = n_c\gamma_c(1-\gamma_c)$. Then

$$\begin{aligned} Var(\hat{w}_c) = Var\left(\frac{n_D^c}{n_D}\right) &\approx \left(\frac{n_c\gamma_c}{\sum_{c=1}^m n_c\gamma_c}\right)^2 \left\{\frac{1-\gamma_c}{n_c\gamma_c} - 2\frac{1-\gamma_c}{\sum_{c=1}^m n_c\gamma_c} + \frac{\sum_{c=1}^m n_c\gamma_c(1-\gamma_c)}{(\sum_{c=1}^m n_c\gamma_c)^2}\right\} \\ &\leq \frac{1-\gamma_c}{n_c\gamma_c} + \frac{\sum_{c=1}^m n_c\gamma_c(1-\gamma_c)}{(\sum_{c=1}^m n_c\gamma_c)^2}. \end{aligned}$$

By Hölder's inequality

$$\sum_{c=1}^m \sqrt{Var(\hat{w}_c)} \leq \sqrt{mA},$$

where

$$A = \sum_{c=1}^m \left\{\frac{1-\gamma_c}{n_c\gamma_c} + \frac{\sum_{c=1}^m n_c\gamma_c(1-\gamma_c)}{(\sum_{c=1}^m n_c\gamma_c)^2}\right\}.$$

We can write

$$A = \sum_{c=1}^m \frac{1-\gamma_c}{n_c\gamma_c} + m\left\{\frac{\sum_{c=1}^m n_c\gamma_c(1-\gamma_c)}{(\sum_{c=1}^m n_c\gamma_c)^2}\right\} \leq \frac{1}{\min_c n_c}\sum_{c=1}^m (1/\gamma_c) + \frac{m}{\sum_{c=1}^m n_c\gamma_c}.$$

Furthermore, we have $1/\gamma_c \leq V < \infty$. Then

$$A \leq \frac{mV}{\min_c n_c} + \frac{m}{\min_c n_c \sum_{c=1}^m \gamma_c} \leq \frac{(m+1)V}{\min_c n_c} \approx \frac{2mV}{\min_c n_c}.$$

Then $\sum_{c=1}^m \sqrt{Var(\hat{w}_c)} \to 0$ if $n_c \to \infty$, $c = 1, ..., m$, and $m \to M$ such that $\frac{\sqrt{\min_c n_c}}{m} \to \infty$. This holds if $n_c \to \infty$, $c = 1, ..., m$, and $m \to M$ such that $\sqrt{n_c}/m \to \infty$. This then gives $P(\sum_{c=1}^m |\hat{w}_c - w_c| > \epsilon) \to 0$, completing the proof. □

Theorem 2. *Suppose* $\hat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}_0$ *as* $n_c \to \infty$, $c = 1, ..., m$, *and* $m \to M$ *such that* $\sqrt{n_c}/m \to \infty$. *Further suppose that conditions (C1)–(C6) hold. Then* $aA\hat{U}C(\hat{\boldsymbol{\theta}}) \xrightarrow{p} aAUC(\boldsymbol{\theta}_0)$ *as* $n_c \to \infty$, $c = 1, ..., m$, *and* $m \to M$ *such that* $\sqrt{n_c}/m \to \infty$.

Proof. We can write this claim as

$$\left|aA\hat{U}C(\hat{\boldsymbol{\theta}}) - aAUC(\boldsymbol{\theta}_0)\right| = \left|aA\hat{U}C(\hat{\boldsymbol{\theta}}) - aAUC(\hat{\boldsymbol{\theta}}) + aAUC(\hat{\boldsymbol{\theta}}) - aAUC(\boldsymbol{\theta}_0)\right| = o_p(1).$$

Then

$$\left|aA\hat{U}C(\hat{\boldsymbol{\theta}}) - aAUC(\boldsymbol{\theta}_0)\right| \leq \left|aA\hat{U}C(\hat{\boldsymbol{\theta}}) - aAUC(\hat{\boldsymbol{\theta}})\right| + \left|aAUC(\hat{\boldsymbol{\theta}}) - aAUC(\boldsymbol{\theta}_0)\right|.$$

By the uniform convergence of $aA\hat{U}C(\boldsymbol{\theta})$ (Theorem 1),

$$\left|aA\hat{U}C(\hat{\boldsymbol{\theta}}) - aAUC(\hat{\boldsymbol{\theta}})\right| \leq \sup_{\boldsymbol{\theta} \in B} \left|aA\hat{U}C(\boldsymbol{\theta}) - aAUC(\boldsymbol{\theta})\right| = o_p(1).$$

Next, we can write

$$\begin{aligned} \left|aAUC(\hat{\boldsymbol{\theta}}) - aAUC(\boldsymbol{\theta}_0)\right| &= \left|\sum_{c=1}^{M} w_c(AUC_c(\hat{\boldsymbol{\theta}}) - AUC_c(\boldsymbol{\theta}_0))\right| \\ &\leq \sum_{c=1}^{M} w_c \left|AUC_c(\hat{\boldsymbol{\theta}}) - AUC_c(\boldsymbol{\theta}_0)\right| \end{aligned}$$

Then we can apply Taylor's theorem to $AUC_c(\hat{\boldsymbol{\theta}}) - AUC_c(\boldsymbol{\theta}_0)$ using condition (C5). This gives

$$AUC_c(\hat{\boldsymbol{\theta}}) - AUC_c(\boldsymbol{\theta}_0) \approx \left(\frac{\partial}{\partial \mathbf{t}} AUC_c(\mathbf{t})\Big|_{\mathbf{t}=\boldsymbol{\theta}_0}\right)^{\top} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0).$$

Then

$$\sum_{c=1}^{M} w_c \left| AUC_c(\hat{\boldsymbol{\theta}}) - AUC_c(\boldsymbol{\theta}_0) \right| \approx \sum_{c=1}^{M} w_c \left| AUC_c'(\boldsymbol{\theta}_0)^\top (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \right|$$

$$\leq \sqrt{p} \times T \times ||\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0|| \times \sum_{c=1}^{M} w_c = o_p(1),$$

by condition (C5), the Cauchy-Schwarz inequality, the convergence of $\hat{\boldsymbol{\theta}}$, the continuous mapping theorem, and the fact that $\sum_{c=1}^{M} w_c = 1$. □

S3 Propositions 1 and 2

Proposition 1. *Suppose* $C \perp L_{\boldsymbol{\theta}}(\boldsymbol{X})|\bar{D}$ *for a combination based on some* $\boldsymbol{\theta}$ *such that* $L_{\boldsymbol{\theta}}(\boldsymbol{X})$ *has common support among cases and controls. Then* $AUC(\boldsymbol{\theta}) = aAUC(\boldsymbol{\theta})$.

This proposition is presented without proof; the result follows directly from Result 6.2 in Pepe.[2]

Proposition 2. *Suppose* $C \perp D$ *and for a combination based on some* $\boldsymbol{\theta}$, $AUC_c(\boldsymbol{\theta}) = AUC^*(\boldsymbol{\theta}),\ \ c = 1, ..., M$ *and* $ROC_c(\boldsymbol{\theta})$ *is concave for* $c = 1, ..., M$. *Then* $AUC_c(\boldsymbol{\theta}) \geq AUC(\boldsymbol{\theta})$, $c = 1, ..., M$.

This proposition is presented without proof, as the result follows directly from Result 6.1 in Pepe,[2] but here we have a linear combination $L_{\boldsymbol{\theta}}(\mathbf{X})$ instead of a single marker.

S4 Center-Specific AUC for Conditionally Bivariate Normal Biomarkers

Claim: If the biomarkers X_1 and X_2 have the conditional distribution given by equation (6), $AUC_c(\boldsymbol{\theta})$ for a generic combinations $\boldsymbol{\theta}$ does not vary with c and the center-specific ROC curves for $\boldsymbol{\theta}^\top \mathbf{X}$ are concave.

Proof. Let X_{1D} denote X_1 for an arbitrary case and $X_{1\bar{D}}$ denote X_1 for an arbitrary control (and define X_{2D} and $X_{2\bar{D}}$ analogously). The $AUC_c(\boldsymbol{\theta})$ for a generic $\boldsymbol{\theta}$ can be written as:

$$\begin{aligned} AUC_c(\boldsymbol{\theta}) &= P(\theta_1 X_{1D} + \theta_2 X_{2D} > \theta_1 X_{1\bar{D}} + \theta_2 X_{2\bar{D}} C = c) \\ &= P\left(\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}^\top \begin{pmatrix} X_{1D} - X_{1\bar{D}} \\ X_{2D} - X_{2\bar{D}} \end{pmatrix} > 0 \middle| C = c \right), \end{aligned}$$

where the distribution of $\left(\begin{matrix} X_{1D} - X_{1\bar{D}} \\ X_{2D} - X_{2\bar{D}} \end{matrix} \middle| C = c \right)$ is constant across centers under the distribution given by equation (7) in Section 5.2. Since $\theta_1 X_{1D} + \theta_2 X_{2D}$ and $\theta_1 X_{1\bar{D}} + \theta_2 X_{2\bar{D}}$ are independently normally distributed (conditional on $C = c$) with

$$\begin{aligned} &(\theta_1 X_{1D} + \theta_2 X_{2D} | C = c) \sim N(\theta_1 \{\mu_{X_1} + f_{X_1}(c)\} + \theta_2 \{\mu_{X_2} + f_{X_2}(c)\}, \theta_1^2 + \theta_2^2 + 2\rho\theta_1\theta_2) \\ &(\theta_1 X_{1\bar{D}} + \theta_2 X_{2\bar{D}} | C = c) \sim N(\theta_1 f_{X_1}(c) + \theta_2 f_{X_2}(c), \theta_1^2 + \theta_2^2 + 2\rho\theta_1\theta_2), \end{aligned}$$

the center-specific AUC is a function of the variances of $(\theta_1 X_{1D} + \theta_2 X_{2D} | C = c)$ and $(\theta_1 X_{1\bar{D}} + \theta_2 X_{2\bar{D}} | C = c)$ and the difference in the means of $(\theta_1 X_{1D} + \theta_2 X_{2D} | C = c)$ and $(\theta_1 X_{1\bar{D}} + \theta_2 X_{2\bar{D}} | C = c)$.[2] The difference in the means is $\theta_1 \mu_{X_1} + \theta_2 \mu_{X_2}$, and the variances also do not depend on center. Thus, $AUC_c(\boldsymbol{\theta})$ does not vary with c. Furthermore, since within each center center, the distribution of the combination in cases and controls is normally distributed with equal variance, the center-specific ROC curves for the combination are concave.[2] □

S5 Additional Simulation Results

S5.1 Ignoring Center

The tables below present the full results for the simulations assessing the impact of ignoring center in development and/or evaluation. In each table, we present the average across simulations of the coefficient estimates and AUCs, the percent bias, and the mean squared error (MSE). The percent bias and MSE for $\hat{\alpha}_1$ and $\hat{\beta}_1$ are relative to β_1, the percent bias and MSE for $\hat{\alpha}_2$ and $\hat{\beta}_2$ are relative to β_2, and the percent bias and MSE for $AUC(\hat{\boldsymbol{\alpha}})$, $AUC_c(\hat{\boldsymbol{\alpha}})$, $AUC(\hat{\boldsymbol{\beta}})$, and $AUC_c(\hat{\boldsymbol{\beta}})$ are relative to $AUC_c(\boldsymbol{\beta})$. We have multiplied the MSE by 10^4. The label "Confounder (+)" refers to the setting of positive correlation between logit(γ_c) and $f(c)$, while "Confounder (-)" refers to the setting of negative correlation between logit(γ_c) and $f(c)$. Although we refer to the differences between the parameter estimates ($\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$) and $\boldsymbol{\beta}$ and between the AUC values ($AUC(\hat{\boldsymbol{\alpha}})$, $AUC_c(\hat{\boldsymbol{\alpha}})$, $AUC(\hat{\boldsymbol{\beta}})$, and $AUC_c(\hat{\boldsymbol{\beta}})$) and $AUC_c(\boldsymbol{\beta})$ as "bias," $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ reflect different population parameters (in general), as do $AUC(\boldsymbol{\alpha})$, $AUC_c(\boldsymbol{\alpha})$, $AUC(\boldsymbol{\beta})$, and $AUC_c(\boldsymbol{\beta})$.

S5.1.1 6 Centers

Table S1: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates and the AUC across 500 simulations when there were 6 centers in the training data and center was either a case mix variable, a calibration variable, a confounder with positive correlation between logit(γ_c) and $f(c)$ ("Confounder (+)") or a confounder with negative correlation between logit(γ_c) and $f(c)$ ("Confounder (–)"). The coefficient estimates ($\hat{\alpha}_1, \hat{\alpha}_2$) and marginal AUC ($AUC(\cdot)$) correspond to ignoring center during construction and evaluation, respectively. The coefficient estimates ($\hat{\beta}_1, \hat{\beta}_2$) and conditional AUC ($AUC_c(\cdot)$) correspond to accounting for center during construction and evaluation, respectively. The MSE is multiplied by 10^4.

	Case Mix			Calibration			Confounder (+)			Confounder (–)		
	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
$\hat{\alpha}_1$	0.116	1.468	50.07	-0.124	-207.952	613.40	0.002	-98.554	211.96	-0.251	-319.351	1429.40
$\hat{\alpha}_2$	0.491	0.625	59.57	0.249	-48.974	625.19	0.374	-23.223	205.83	0.125	-74.306	1408.40
$\hat{\beta}_1$	0.116	1.660	61.30	0.117	2.105	45.55	0.118	3.308	59.96	0.112	-2.252	62.56
$\hat{\beta}_2$	0.492	0.783	71.89	0.490	0.483	49.31	0.490	0.523	62.94	0.490	0.459	61.89
$AUC(\hat{\boldsymbol{\alpha}})$	0.651	-0.133	0.04	0.577	-11.618	59.05	0.707	8.360	66.59	0.577	-11.495	118.38
$AUC_c(\hat{\boldsymbol{\alpha}})$	0.651	-0.130	0.04	0.612	-6.246	22.42	0.646	-0.972	1.44	0.479	-26.582	349.41
$AUC(\hat{\boldsymbol{\beta}})$	0.651	-0.163	0.06	0.574	-12.036	64.59	0.708	8.544	68.43	0.432	-33.804	556.73
$AUC_c(\hat{\boldsymbol{\beta}})$	0.651	-0.162	0.06	0.652	-0.119	0.03	0.651	-0.162	0.05	0.651	-0.164	0.05

S5.1.2 500 Centers

Table S2: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates and the AUC across 500 simulations when there were 500 centers in the training data and center was either a case mix variable, a calibration variable, a confounder with positive correlation between logit(γ_c) and $f(c)$ ("Confounder (+)") or a confounder with negative correlation between logit(γ_c) and $f(c)$ ("Confounder (–)"). The coefficient estimates ($\hat{\alpha}_1, \hat{\alpha}_2$) and marginal AUC ($AUC(\cdot)$) correspond to ignoring center during construction and evaluation, respectively. The coefficient estimates ($\hat{\beta}_1, \hat{\beta}_2$) and conditional AUC ($AUC_c(\cdot)$) correspond to accounting for center during construction and evaluation, respectively. The MSE is multiplied by 10^4.

	Case Mix			Calibration			Confounder (+)			Confounder (–)		
	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
$\hat{\alpha}_1$	0.114	-0.713	5.71	-0.147	-228.472	688.20	-0.014	-112.174	169.94	-0.275	-340.089	1519.49
$\hat{\alpha}_2$	0.488	-0.028	5.80	0.226	-53.703	690.71	0.363	-25.627	161.58	0.097	-80.021	1527.88
$\hat{\beta}_1$	0.114	-0.808	7.41	0.115	0.900	7.06	0.113	-1.420	6.99	0.114	-0.380	7.02
$\hat{\beta}_2$	0.487	-0.071	7.44	0.488	0.046	6.75	0.490	0.390	7.60	0.488	0.094	7.80
$AUC(\hat{\boldsymbol{\alpha}})$	0.652	-0.026	0.02	0.574	-12.016	61.47	0.719	10.293	45.54	0.624	-4.402	8.91
$AUC_c(\hat{\boldsymbol{\alpha}})$	0.652	-0.026	0.02	0.599	-8.170	28.81	0.649	-0.531	0.17	0.447	-31.473	422.54
$AUC(\hat{\boldsymbol{\beta}})$	0.652	-0.031	0.02	0.559	-14.290	86.94	0.719	10.169	44.47	0.395	-39.418	662.03
$AUC_c(\hat{\boldsymbol{\beta}})$	0.652	-0.031	0.02	0.652	-0.042	0.02	0.652	-0.015	0.02	0.652	-0.026	0.02

S5.2 RILR vs. FILR

The tables below present the full results for the simulations comparing RILR and FILR for constructing combinations. The results are separated by the role of center (i.e., case mix variable, calibration variable, or confounder with negative, no, or positive correlation between $f(c)$ and $\text{logit}(\gamma_c)$) and the distribution of $\text{logit}(\gamma_c)$ and/or $f(c)$, F (i.e., normal, Gumbel, Laplace, or uniform). In the first table we report the mean, percent bias, and mean squared error (MSE) for the coefficient estimates based on RILR ($\hat{\boldsymbol{\tau}} = (\hat{\tau}_1, \hat{\tau}_2)$) and FILR ($\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \hat{\beta}_2)$) relative to the true coefficients ($\boldsymbol{\beta} = (\beta_1, \beta_2)$), as well as the average, percent bias and MSE for the conditional AUCs based on the estimated coefficients ($AUC_c(\hat{\boldsymbol{\tau}})$ and $AUC_c(\hat{\boldsymbol{\beta}})$, respectively) relative to $AUC_c(\boldsymbol{\beta})$. In the second table we report the average, percent bias and MSE for the overall (fixed) intercept estimate provided by RILR ($\hat{\tau}_0$). In the setting where center is a calibration variable, asterisks indicate the results for $\gamma_c = 0.1$; the other calibration variable results are for $\gamma_c = 0.5$. We have multiplied the MSE by 10^4. As noted above, the "bias" is calculated as the difference between the parameter estimates and $\boldsymbol{\beta}$ and between the AUC values and $AUC_c(\boldsymbol{\beta})$; it is not exactly accurate to call these differences biases, as they arise because (in the case of $\hat{\boldsymbol{\tau}}$ and $AUC_c(\hat{\boldsymbol{\tau}})$) they correspond to different population parameters (in general).

S5.2.1 Case Mix

Normal

Table S3: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression ($\hat{\tau}_1, \hat{\tau}_2$) and fixed intercept logistic regression ($\hat{\beta}_1, \hat{\beta}_2$) and center-specific AUC ($AUC_c(\cdot)$) across 500 simulations when center was a case mix variable and logit(γ_c) was normally distributed. The number of centers in the training data, m, and the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{\gamma_c}$), denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
								$m = 6$										
1	0.114	-0.07	60.07	0.114	-0.08	60.71	0.494	1.19	58.32	0.494	1.23	59.01	0.651	-0.18	0.39	0.651	-0.19	0.39
0.5	0.115	0.54	53.28	0.115	0.53	53.95	0.493	1.15	55.62	0.493	1.18	56.25	0.651	-0.14	0.36	0.651	-0.15	0.37
0.5/1.5	0.116	1.09	60.76	0.116	1.18	61.39	0.490	0.47	61.84	0.490	0.51	62.67	0.651	-0.18	0.38	0.651	-0.18	0.38
3	0.117	2.23	75.12	0.117	2.32	76.06	0.492	0.79	90.99	0.492	0.82	92.23	0.651	-0.16	0.79	0.651	-0.16	0.79
1/5	0.112	-2.12	75.17	0.112	-2.08	75.93	0.485	-0.60	75.80	0.485	-0.60	76.69	0.651	-0.23	0.44	0.651	-0.23	0.44
								$m = 500$										
1	0.119	3.79	7.87	0.114	-0.53	7.75	0.510	4.59	13.01	0.489	0.23	8.08	0.652	-0.06	0.42	0.652	-0.06	0.42
0.5	0.119	3.57	7.04	0.115	0.07	7.07	0.505	3.61	9.94	0.488	-0.03	6.91	0.653	0.04	0.34	0.653	0.04	0.34
0.5/1.5	0.119	4.30	6.44	0.114	-0.02	6.18	0.510	4.56	11.85	0.489	0.23	6.88	0.652	0.02	0.33	0.652	0.01	0.33
3	0.120	5.24	9.72	0.115	0.26	9.56	0.511	4.82	14.82	0.487	-0.16	9.41	0.652	0.01	0.84	0.652	0.00	0.84
1/5	0.120	5.10	9.10	0.115	0.11	8.68	0.514	5.44	15.82	0.490	0.38	8.64	0.653	0.04	0.37	0.653	0.03	0.37

Table S4: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a case mix variable and logit(γ_c) was normally distributed. The number of centers in the training data, m, and the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{\gamma_c}$), denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	Mean	Bias	MSE
	$m = 6$		
1	-0.149	-3.00	1777.65
0.5	-0.161	4.72	1021.73
0.5/1.5	-0.189	23.10	1701.55
3	-0.173	12.56	4994.47
1/5	-0.079	-48.24	5019.95
	$m = 500$		
1	-0.158	2.92	25.53
0.5	-0.162	5.71	15.04
0.5/1.5	-0.162	5.39	27.89
3	-0.163	6.27	65.01
1/5	-0.167	8.93	58.85

Gumbel

Table S5: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression ($\hat{\tau}_1, \hat{\tau}_2$) and fixed intercept logistic regression ($\hat{\beta}_1, \hat{\beta}_2$) and center-specific AUC ($AUC_c(\cdot)$) across 500 simulations when center was a case mix variable and logit(γ_c) had a Gumbel distribution. The number of centers in the training data, m, and the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{\gamma_c}$), denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
									$m = 6$									
1	0.111	-2.69	59.00	0.111	-2.70	59.68	0.492	0.87	68.51	0.492	0.88	69.21	0.651	-0.22	0.47	0.651	-0.22	0.47
0.5	0.120	4.71	57.30	0.120	4.80	57.91	0.492	0.84	56.23	0.492	0.85	56.84	0.651	-0.15	0.37	0.651	-0.15	0.37
0.5/1.5	0.118	3.51	53.66	0.118	3.55	54.29	0.493	1.16	55.68	0.494	1.19	56.28	0.651	-0.16	0.42	0.651	-0.16	0.42
3	0.120	4.53	72.70	0.120	4.63	73.66	0.490	0.45	74.68	0.490	0.48	75.59	0.652	-0.04	0.86	0.652	-0.05	0.86
1/5	0.111	-3.25	66.01	0.111	-3.28	66.65	0.497	1.98	69.62	0.497	2.00	70.39	0.651	-0.20	0.59	0.651	-0.20	0.59
									$m = 500$									
1	0.121	5.94	7.34	0.116	1.62	6.89	0.509	4.38	12.25	0.488	0.07	7.64	0.652	-0.00	0.47	0.652	-0.00	0.47
0.5	0.120	4.99	7.29	0.116	1.45	7.21	0.504	3.35	9.73	0.487	-0.22	7.26	0.652	-0.03	0.34	0.652	-0.03	0.34
0.5/1.5	0.121	5.39	6.75	0.116	1.05	6.35	0.507	4.00	10.24	0.486	-0.26	6.48	0.653	0.04	0.34	0.653	0.04	0.33
3	0.120	4.70	8.30	0.114	-0.21	8.15	0.512	4.95	13.81	0.488	-0.02	8.11	0.652	-0.07	0.74	0.652	-0.08	0.74
1/5	0.118	3.49	8.06	0.113	-1.48	7.94	0.515	5.51	15.56	0.490	0.49	8.40	0.652	-0.06	0.39	0.652	-0.06	0.39

Table S6: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a case mix variable and logit(γ_c) had a Gumbel distribution. The number of centers in the training data, m, and the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{\gamma_c}$), denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	Mean	Bias	MSE
	$m = 6$		
1	-0.124	-19.47	1864.27
0.5	-0.172	12.02	822.96
0.5/1.5	-0.164	6.96	1763.39
3	-0.113	-26.10	5151.42
1/5	-0.104	-32.38	5129.53
	$m = 500$		
1	-0.195	26.84	38.91
0.5	-0.174	13.28	18.73
0.5/1.5	-0.192	25.32	35.47
3	-0.260	69.38	175.98
1/5	-0.251	63.91	146.37

Laplace

Table S7: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression $(\hat{\tau}_1, \hat{\tau}_2)$ and fixed intercept logistic regression $(\hat{\beta}_1, \hat{\beta}_2)$ and center-specific AUC $(AUC_c(\cdot))$ across 500 simulations when center was a case mix variable and logit(γ_c) had a Laplace distribution. The number of centers in the training data, m, and the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{\gamma_c}$), denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
									$m = 6$									
1	0.113	-1.56	62.51	0.113	-1.62	63.17	0.491	0.65	58.82	0.491	0.68	59.48	0.651	-0.14	0.47	0.651	-0.14	0.48
0.5	0.117	2.25	55.74	0.117	2.29	56.41	0.491	0.65	56.96	0.491	0.67	57.47	0.651	-0.13	0.37	0.651	-0.13	0.37
0.5/1.5	0.115	0.69	62.09	0.115	0.71	62.71	0.494	1.33	57.52	0.494	1.33	58.00	0.651	-0.22	0.42	0.651	-0.22	0.42
3	0.112	-1.94	75.19	0.112	-1.97	76.07	0.493	1.03	68.33	0.493	1.04	69.13	0.652	-0.12	3.28	0.652	-0.12	3.28
1/5	0.117	2.32	64.94	0.117	2.36	65.50	0.493	1.07	72.65	0.493	1.07	73.23	0.651	-0.19	0.43	0.651	-0.19	0.43
									$m = 500$									
1	0.121	5.77	6.84	0.116	1.47	6.51	0.508	4.14	11.39	0.487	-0.14	7.38	0.652	-0.08	0.44	0.652	-0.08	0.44
0.5	0.119	4.04	6.30	0.115	0.26	6.10	0.505	3.51	9.37	0.488	-0.01	6.45	0.652	0.02	0.33	0.652	0.02	0.33
0.5/1.5	0.117	2.64	7.14	0.112	-1.70	7.13	0.508	4.14	11.18	0.487	-0.07	7.36	0.652	0.00	0.35	0.652	-0.00	0.35
3	0.122	6.71	9.88	0.116	1.54	9.42	0.511	4.78	15.05	0.486	-0.29	9.70	0.653	0.13	2.22	0.653	0.12	2.22
1/5	0.119	3.92	8.15	0.113	-1.06	8.03	0.513	5.27	14.60	0.489	0.20	8.02	0.652	-0.10	0.39	0.652	-0.10	0.39

Table S8: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a case mix variable and logit(γ_c) had a Laplace distribution. The number of centers in the training data, m, and the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{\gamma_c}$), denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	Mean	Bias	MSE
	$m = 6$		
1	-0.191	24.52	1881.44
0.5	-0.160	4.31	938.70
0.5/1.5	-0.155	0.82	1713.69
3	-0.184	20.28	4805.53
1/5	-0.234	52.49	5052.67
	$m = 500$		
1	-0.159	3.71	20.97
0.5	-0.158	2.75	14.74
0.5/1.5	-0.163	5.97	21.87
3	-0.155	1.34	49.87
1/5	-0.157	2.59	48.05

Uniform

Table S9: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression $(\hat{\tau}_1, \hat{\tau}_2)$ and fixed intercept logistic regression $(\hat{\beta}_1, \hat{\beta}_2)$ and center-specific AUC $(AUC_c(\cdot))$ across 500 simulations when center was a case mix variable and logit(γ_c) had a uniform distribution. The number of centers in the training data, m, and the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{\gamma_c}$), denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
									$m = 6$									
1	0.117	2.55	63.17	0.117	2.62	63.93	0.491	0.75	63.06	0.492	0.79	63.74	0.651	-0.14	0.44	0.651	-0.14	0.44
0.5	0.117	2.17	55.26	0.117	2.23	55.68	0.488	0.15	58.61	0.489	0.17	59.00	0.652	-0.12	0.36	0.652	-0.12	0.36
0.5/1.5	0.116	1.33	59.55	0.116	1.35	60.25	0.496	1.70	62.01	0.496	1.74	62.71	0.652	-0.10	0.35	0.652	-0.10	0.35
3	0.112	-2.50	79.76	0.112	-2.45	81.06	0.498	2.13	77.55	0.499	2.25	78.84	0.651	-0.27	0.76	0.651	-0.28	0.76
1/5	0.124	8.61	78.28	0.124	8.70	79.40	0.490	0.52	80.29	0.490	0.56	81.50	0.651	-0.26	0.46	0.651	-0.26	0.46
									$m = 500$									
1	0.120	4.91	7.14	0.115	0.75	6.91	0.509	4.42	12.14	0.488	-0.00	7.55	0.652	-0.12	0.39	0.652	-0.12	0.39
0.5	0.119	3.94	6.69	0.115	0.27	6.68	0.506	3.70	10.66	0.487	-0.04	7.54	0.652	-0.01	0.30	0.652	-0.01	0.30
0.5/1.5	0.118	3.40	7.14	0.113	-1.09	7.04	0.510	4.50	12.78	0.488	0.13	7.81	0.653	0.03	0.33	0.653	0.03	0.33
3	0.117	2.10	10.38	0.111	-2.75	10.65	0.514	5.39	16.39	0.491	0.61	9.66	0.652	-0.09	0.69	0.652	-0.10	0.69
1/5	0.119	4.29	9.84	0.113	-0.93	9.72	0.513	5.16	15.87	0.489	0.19	9.63	0.652	-0.01	0.39	0.652	-0.01	0.39

Table S10: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a case mix variable and logit(γ_c) had a uniform distribution. The number of centers in the training data, m, and the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{\gamma_c}$), denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	Mean	Bias	MSE
	$m = 6$		
1	-0.157	2.38	1665.04
0.5	-0.160	4.31	833.86
0.5/1.5	-0.141	-7.92	1734.34
3	-0.184	19.84	4863.75
1/5	-0.163	6.16	5291.33
	$m = 500$		
1	-0.165	7.83	29.53
0.5	-0.161	5.19	17.88
0.5/1.5	-0.154	0.56	28.13
3	-0.157	2.46	91.62
1/5	-0.164	6.62	73.88

S5.2.2 Calibration

Normal

Table S11: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression ($\hat{\tau}_1, \hat{\tau}_2$) and fixed intercept logistic regression ($\hat{\beta}_1, \hat{\beta}_2$) and center-specific AUC ($AUC_c(\cdot)$) across 500 simulations when center was a calibration variable and $f(c)$ was normally distributed. The number of centers in the training data, m, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{f(c)}$), denoted by "a/b" in the $\sigma^2_{f(c)}$ column; in these scenarios, half of the centers had $\sigma^2_{f(c)} = a$ and half had $\sigma^2_{f(c)} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
								$m = 6$										
5	0.092	-19.56	51.35	0.118	3.29	44.79	0.463	-4.98	62.24	0.490	0.40	54.41	0.651	-0.19	0.29	0.651	-0.16	0.28
1	0.092	-19.63	55.56	0.117	2.13	48.57	0.467	-4.34	54.61	0.491	0.74	48.48	0.651	-0.19	0.32	0.651	-0.16	0.30
2/8	0.088	-23.30	58.45	0.114	-0.61	49.57	0.468	-4.10	60.69	0.494	1.27	55.32	0.651	-0.16	0.33	0.652	-0.12	0.33
*5	0.018	-84.31	281.73	0.118	2.83	131.50	0.392	-19.58	281.08	0.492	0.81	134.10	0.640	-1.82	6.59	0.650	-0.36	1.06
*2/8	0.029	-74.61	246.80	0.124	8.61	122.84	0.395	-18.95	282.48	0.491	0.77	137.85	0.642	-1.62	5.99	0.651	-0.25	0.85
								$m = 500$										
5	-0.142	-223.75	660.10	0.116	1.06	6.10	0.233	-52.13	651.20	0.488	0.14	6.63	0.605	-7.32	23.56	0.652	-0.07	0.29
1	-0.030	-126.22	214.42	0.115	0.78	6.46	0.348	-28.60	200.80	0.488	0.11	6.46	0.647	-0.78	0.63	0.652	-0.02	0.29
2/8	-0.142	-224.31	663.65	0.115	0.29	6.73	0.235	-51.88	644.59	0.489	0.31	6.12	0.605	-7.33	23.73	0.652	0.00	0.28
*5	-0.147	-228.13	693.12	0.114	-0.24	16.19	0.231	-52.60	669.62	0.490	0.41	15.82	0.602	-7.78	27.45	0.652	-0.06	0.84
*2/8	-0.145	-226.72	684.68	0.117	1.86	16.86	0.230	-52.85	675.66	0.490	0.45	16.25	0.602	-7.67	27.03	0.652	-0.01	0.76

Table S12: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a calibration variable and $f(c)$ was normally distributed. The number of centers in the training data, m, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{f(c)}$), denoted by "a/b" in the $\sigma^2_{f(c)}$ column; in these scenarios, half of the centers had $\sigma^2_{f(c)} = a$ and half had $\sigma^2_{f(c)} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{f(c)}$	Mean	Bias	MSE
5	-0.137	-10.52	2840.66
1	-0.141	-8.16	543.63
2/8	-0.112	-27.27	2778.08
*5	-2.313	-1.59	1816.01
*2/8	-2.336	-0.62	1740.03
	$m = 500$		
5	-0.038	-75.24	138.40
1	-0.089	-41.89	48.34
2/8	-0.039	-74.77	136.83
*5	-2.249	-4.31	116.82
*2/8	-2.251	-4.22	113.70

Gumbel

Table S13: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression $(\hat{\tau}_1, \hat{\tau}_2)$ and fixed intercept logistic regression $(\hat{\beta}_1, \hat{\beta}_2)$ and center-specific AUC $(AUC_c(\cdot))$ across 500 simulations when center was a calibration variable and $f(c)$ had a Gumbel distribution. The number of centers in the training data, m, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{f(c)}$), denoted by "a/b" in the $\sigma^2_{f(c)}$ column; in these scenarios, half of the centers had $\sigma^2_{f(c)} = a$ and half had $\sigma^2_{f(c)} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
									$m = 6$									
5	0.082	-27.99	60.42	0.109	-5.06	48.57	0.469	-3.86	54.68	0.495	1.53	50.20	0.651	-0.18	0.34	0.651	-0.13	0.32
1	0.093	-18.44	57.45	0.118	2.86	50.76	0.466	-4.47	57.73	0.490	0.51	51.57	0.651	-0.17	0.31	0.651	-0.14	0.30
2/8	0.089	-22.09	52.15	0.115	0.58	43.91	0.468	-3.98	58.50	0.494	1.38	52.91	0.651	-0.17	0.31	0.651	-0.13	0.30
*5	0.020	-82.71	284.51	0.117	2.42	131.37	0.399	-18.09	253.74	0.496	1.77	120.85	0.639	-2.00	8.04	0.650	-0.39	0.99
*2/8	0.021	-81.87	249.45	0.112	-2.31	118.70	0.411	-15.66	237.42	0.503	3.12	135.69	0.642	-1.58	5.21	0.649	-0.45	1.05
									$m = 500$									
5	-0.141	-223.20	657.15	0.115	0.79	5.93	0.233	-52.16	652.07	0.486	-0.25	7.07	0.605	-7.28	23.30	0.652	-0.05	0.30
1	-0.031	-127.22	217.23	0.114	-0.38	5.50	0.350	-28.33	196.60	0.489	0.33	5.62	0.647	-0.83	0.64	0.652	-0.04	0.31
2/8	-0.142	-223.66	660.11	0.115	0.14	7.07	0.235	-51.81	643.94	0.489	0.24	6.68	0.605	-7.25	23.20	0.652	-0.03	0.26
*5	-0.149	-229.91	704.29	0.114	-0.40	16.52	0.229	-53.03	681.50	0.491	0.64	18.40	0.599	-8.12	29.95	0.652	-0.06	0.91
*2/8	-0.149	-229.97	706.01	0.113	-1.05	16.49	0.228	-53.30	689.28	0.488	-0.02	18.59	0.600	-8.10	29.72	0.652	-0.03	0.79

Table S14: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a calibration variable and $f(c)$ had a Gumbel distribution. The number of centers in the training data, m, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{f(c)}$), denoted by "a/b" in the $\sigma^2_{f(c)}$ column; in these scenarios, half of the centers had $\sigma^2_{f(c)} = a$ and half had $\sigma^2_{f(c)} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{f(c)}$	Mean	Bias	MSE
	$m = 6$		
5	-0.154	0.20	2584.85
1	-0.137	-10.48	510.19
2/8	-0.155	0.90	2809.27
*5	-2.329	-0.94	1926.27
*2/8	-2.333	-0.76	2183.85
	$m = 500$		
5	-0.038	-74.98	137.05
1	-0.089	-42.21	47.75
2/8	-0.038	-75.25	137.85
*5	-2.249	-4.34	118.27
*2/8	-2.247	-4.42	122.61

Laplace

Table S15: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression ($\hat{\tau}_1, \hat{\tau}_2$) and fixed intercept logistic regression ($\hat{\beta}_1, \hat{\beta}_2$) and center-specific AUC ($AUC_c(\cdot)$) across 500 simulations when center was a calibration variable and $f(c)$ had a Laplace distribution The number of centers in the training data, m, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{f(c)}$), denoted by "a/b" in the $\sigma^2_{f(c)}$ column; in these scenarios, half of the centers had $\sigma^2_{f(c)} = a$ and half had $\sigma^2_{f(c)} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
								$m = 6$										
5	0.088	-23.50	63.25	0.113	-0.83	54.19	0.467	-4.27	56.90	0.493	1.05	51.56	0.651	-0.24	0.32	0.651	-0.20	0.31
1	0.100	-12.54	51.86	0.124	8.10	48.92	0.464	-4.84	59.25	0.488	0.02	52.40	0.651	-0.15	0.34	0.651	-0.13	0.32
2/8	0.088	-23.17	58.68	0.114	-0.51	49.62	0.462	-5.20	58.70	0.488	0.13	50.48	0.651	-0.16	0.34	0.652	-0.12	0.31
*5	0.035	-69.58	222.54	0.127	10.77	118.43	0.394	-19.15	256.50	0.486	-0.32	117.59	0.643	-1.46	5.05	0.650	-0.30	0.88
*2/8	0.015	-86.96	279.64	0.109	-4.85	127.07	0.400	-18.01	241.53	0.494	1.22	117.33	0.640	-1.83	6.99	0.650	-0.34	1.03
								$m = 500$										
5	-0.141	-223.25	657.19	0.115	0.43	5.90	0.235	-51.87	644.43	0.486	-0.26	6.00	0.605	-7.29	23.44	0.652	-0.07	0.31
1	-0.029	-125.69	212.47	0.115	0.48	6.01	0.349	-28.41	197.90	0.488	-0.03	6.11	0.647	-0.82	0.64	0.652	-0.08	0.31
2/8	-0.142	-224.06	661.70	0.114	0.04	5.91	0.236	-51.70	640.67	0.487	-0.05	6.14	0.605	-7.28	23.41	0.652	-0.06	0.29
*5	-0.145	-226.32	682.92	0.116	1.60	16.25	0.230	-52.93	678.92	0.489	0.32	17.64	0.602	-7.72	27.45	0.652	-0.13	0.92
*2/8	-0.147	-228.05	692.73	0.113	-1.10	15.90	0.230	-52.83	675.47	0.488	-0.01	16.08	0.601	-7.84	28.28	0.652	-0.01	0.82

Table S16: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a calibration variable and $f(c)$ had a Laplace distribution. The number of centers in the training data, m, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{f(c)}$), denoted by "a/b" in the $\sigma^2_{f(c)}$ column; in these scenarios, half of the centers had $\sigma^2_{f(c)} = a$ and half had $\sigma^2_{f(c)} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{f(c)}$	Mean	Bias	MSE
	$m = 6$		
5	-0.125	-18.42	2795.42
1	-0.140	-8.70	533.83
2/8	-0.197	28.33	2448.45
*5	-2.337	-0.56	2059.07
*2/8	-2.347	-0.14	1702.48
	$m = 500$		
5	-0.040	-74.17	134.00
1	-0.089	-42.29	48.77
2/8	-0.039	-74.67	136.63
*5	-2.250	-4.26	115.70
*2/8	-2.252	-4.19	111.36

Uniform

Table S17: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression ($\hat{\tau}_1, \hat{\tau}_2$) and fixed intercept logistic regression ($\hat{\beta}_1, \hat{\beta}_2$) and center-specific AUC ($AUC_c(\cdot)$) across 500 simulations when center was a calibration variable and $f(c)$ had a uniform distribution. The number of centers in the training data, m, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{f(c)}$), denoted by "a/b" in the $\sigma^2_{f(c)}$ column; in these scenarios, half of the centers had $\sigma^2_{f(c)} = a$ and half had $\sigma^2_{f(c)} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
$m = 6$																		
5	0.092	-19.99	56.46	0.118	2.83	49.16	0.464	-4.88	56.30	0.490	0.50	49.05	0.652	-0.08	0.35	0.652	-0.05	0.34
1	0.089	-22.35	57.74	0.114	-0.50	49.41	0.470	-3.72	55.11	0.494	1.38	50.84	0.651	-0.15	0.31	0.652	-0.12	0.30
2/8	0.092	-19.84	60.48	0.118	2.96	52.94	0.469	-3.83	58.21	0.495	1.56	53.25	0.651	-0.16	0.33	0.652	-0.12	0.31
*5	0.021	-81.58	289.01	0.123	7.31	134.23	0.388	-20.51	288.44	0.489	0.24	130.62	0.639	-2.00	7.42	0.650	-0.36	0.93
*2/8	0.018	-84.47	304.82	0.119	3.99	143.28	0.385	-21.15	303.48	0.486	-0.38	141.05	0.639	-2.12	7.83	0.650	-0.41	1.04
$m = 500$																		
5	-0.142	-224.17	662.29	0.114	-0.39	6.09	0.233	-52.32	655.66	0.486	-0.31	6.00	0.605	-7.32	23.48	0.652	0.00	0.31
1	-0.030	-126.31	215.05	0.116	1.23	6.79	0.348	-28.69	201.73	0.487	-0.13	6.83	0.647	-0.75	0.60	0.652	-0.01	0.31
2/8	-0.142	-223.81	660.16	0.116	0.94	6.16	0.233	-52.16	651.70	0.487	-0.08	6.56	0.605	-7.33	23.66	0.652	-0.04	0.30
*5	-0.144	-225.68	677.62	0.116	1.79	15.62	0.229	-52.98	678.44	0.489	0.18	14.81	0.602	-7.68	26.94	0.652	-0.00	0.85
*2/8	-0.143	-225.34	677.58	0.117	1.87	16.51	0.226	-53.57	694.59	0.485	-0.50	15.78	0.602	-7.71	27.30	0.653	0.04	0.81

Table S18: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a calibration variable and $f(c)$ had a uniform distribution. The number of centers in the training data, m, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included heteroscedasticity (i.e., non-constant $\sigma^2_{f(c)}$), denoted by "a/b" in the $\sigma^2_{f(c)}$ column; in these scenarios, half of the centers had $\sigma^2_{f(c)} = a$ and half had $\sigma^2_{f(c)} = b$. The MSE is multiplied by 10^4.

$\sigma^2_{f(c)}$	Mean	Bias	MSE
	$m = 6$		
5	-0.133	-13.42	2975.43
1	-0.136	-11.59	600.93
2/8	-0.165	7.44	2669.35
*5	-2.311	-1.68	1848.61
*2/8	-2.363	0.53	1683.09
	$m = 500$		
5	-0.040	-74.02	134.03
1	-0.089	-42.29	48.48
2/8	-0.039	-74.41	136.09
*5	-2.249	-4.31	116.14
*2/8	-2.251	-4.22	115.37

S5.2.3 Confounding (Negative Correlation)

Normal

Table S19: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression ($\hat{\tau}_1, \hat{\tau}_2$) and fixed intercept logistic regression ($\hat{\beta}_1, \hat{\beta}_2$) and center-specific AUC ($AUC_c(\cdot)$) across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was normal with negative correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
$m = 6$																			
1	5	0.104	-9.10	59.55	0.122	6.21	59.00	0.470	-3.55	64.46	0.488	0.04	61.27	0.650	-0.29	0.49	0.651	-0.27	0.48
3	5	0.100	-12.51	72.89	0.115	0.63	72.04	0.474	-2.82	81.19	0.489	0.29	80.14	0.651	-0.27	0.82	0.651	-0.25	0.80
0.5/1.5	2/8	0.099	-13.83	63.43	0.116	1.63	61.18	0.476	-2.50	61.04	0.493	1.12	59.41	0.651	-0.17	0.39	0.651	-0.15	0.38
1/5	2/8	0.099	-13.87	78.82	0.113	-1.64	75.69	0.478	-2.08	82.39	0.492	0.82	82.70	0.651	-0.21	0.46	0.651	-0.19	0.44
$m = 500$																			
1	5	-0.073	-163.61	358.57	0.116	1.03	7.76	0.316	-35.31	305.30	0.486	-0.31	8.10	0.641	-1.78	2.00	0.653	0.09	0.36
3	5	-0.027	-123.83	210.29	0.116	1.46	8.82	0.362	-25.68	165.28	0.488	-0.01	8.93	0.648	-0.67	1.15	0.653	0.08	0.97
0.5/1.5	2/8	-0.076	-166.22	370.17	0.114	-0.73	7.47	0.315	-35.46	307.58	0.487	-0.18	8.16	0.640	-1.96	2.28	0.652	-0.02	0.38
1/5	2/8	-0.031	-127.19	219.76	0.114	-0.26	7.57	0.362	-25.71	166.18	0.489	0.18	9.00	0.647	-0.80	0.76	0.652	-0.03	0.40

Table S20: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was normal with negative correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	Mean	Bias	MSE
		$m = 6$		
1	5	-0.132	-13.80	6332.20
3	5	-0.154	0.44	11359.50
0.5/1.5	2/8	-0.118	-22.99	6772.60
1/5	2/8	-0.241	57.17	12645.23
		$m = 500$		
1	5	-0.074	-52.02	105.60
3	5	-0.087	-43.09	148.56
0.5/1.5	2/8	-0.073	-52.31	106.10
1/5	2/8	-0.095	-37.91	125.76

Gumbel

Table S21: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression $(\hat{\tau}_1, \hat{\tau}_2)$ and fixed intercept logistic regression $(\hat{\beta}_1, \hat{\beta}_2)$ and center-specific AUC $(AUC_c(\cdot))$ across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was Gumbel with negative correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
								$m = 6$											
1	5	0.101	-11.38	55.21	0.119	3.94	53.84	0.465	-4.56	68.62	0.483	-0.95	63.20	0.651	-0.16	0.44	0.651	-0.15	0.44
3	5	0.096	-16.11	80.01	0.111	-2.75	77.82	0.477	-2.16	85.17	0.493	0.99	83.95	0.650	-0.33	1.06	0.650	-0.31	1.04
0.5/1.5	2/8	0.097	-15.38	58.28	0.114	-0.26	54.60	0.472	-3.15	57.94	0.490	0.42	55.29	0.651	-0.19	0.37	0.651	-0.17	0.36
1/5	2/8	0.101	-11.63	78.41	0.114	-0.05	77.27	0.475	-2.70	87.44	0.488	0.09	86.08	0.652	-0.10	0.46	0.652	-0.09	0.45
								$m = 500$											
1	5	-0.074	-165.03	364.51	0.116	1.10	7.28	0.314	-35.66	309.95	0.486	-0.36	6.97	0.640	-1.96	2.28	0.652	-0.03	0.40
3	5	-0.029	-124.92	212.76	0.117	2.61	8.05	0.361	-26.03	171.06	0.488	0.08	9.62	0.647	-0.76	1.41	0.652	-0.03	1.03
0.5/1.5	2/8	-0.079	-168.87	382.07	0.113	-1.46	7.27	0.315	-35.47	307.60	0.488	0.14	7.37	0.638	-2.14	2.66	0.652	-0.05	0.35
1/5	2/8	-0.033	-128.50	225.88	0.113	-0.92	8.46	0.361	-26.07	170.54	0.488	0.07	8.78	0.647	-0.80	0.86	0.652	0.02	0.48

Table S22: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was Gumbel with negative correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	Mean	Bias	MSE
		$m = 6$		
1	5	-0.165	7.86	6362.19
3	5	-0.127	-16.90	11817.45
0.5/1.5	2/8	-0.106	-31.06	6119.61
1/5	2/8	-0.197	28.42	11399.69
		$m = 500$		
1	5	-0.101	-33.93	65.83
3	5	-0.177	15.49	97.39
0.5/1.5	2/8	-0.099	-35.39	66.38
1/5	2/8	-0.165	7.45	86.90

Laplace

Table S23: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression ($\hat{\tau}_1, \hat{\tau}_2$) and fixed intercept logistic regression ($\hat{\beta}_1, \hat{\beta}_2$) and center-specific AUC ($AUC_c(\cdot)$) across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was Laplace with negative correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
								$m = 6$											
1	5	0.103	-9.61	59.75	0.120	5.22	57.72	0.475	-2.55	62.01	0.492	0.92	60.10	0.651	-0.22	0.42	0.651	-0.20	0.41
3	5	0.097	-15.11	75.58	0.111	-2.81	72.48	0.477	-2.13	68.59	0.492	0.79	69.47	0.651	-0.18	1.82	0.651	-0.16	1.83
0.5/1.5	2/8	0.103	-9.98	61.40	0.120	4.54	60.46	0.471	-3.49	66.32	0.487	-0.11	63.27	0.651	-0.18	0.35	0.651	-0.16	0.35
1/5	2/8	0.097	-15.19	73.53	0.111	-3.25	71.59	0.483	-1.01	64.44	0.497	1.81	65.46	0.651	-0.18	0.42	0.651	-0.16	0.41
								$m = 500$											
1	5	-0.080	-170.20	387.24	0.114	-0.74	7.27	0.311	-36.15	319.94	0.488	-0.02	7.26	0.638	-2.22	2.83	0.652	-0.10	0.41
3	5	-0.030	-126.32	217.87	0.116	1.62	7.79	0.360	-26.16	171.03	0.488	0.10	8.03	0.648	-0.66	1.23	0.653	0.10	0.90
0.5/1.5	2/8	-0.081	-170.79	389.89	0.115	0.45	6.38	0.310	-36.49	324.89	0.488	0.01	7.02	0.638	-2.15	2.64	0.653	0.03	0.35
1/5	2/8	-0.038	-133.37	242.00	0.113	-1.34	8.05	0.355	-27.22	185.18	0.487	-0.15	8.54	0.646	-0.94	0.92	0.652	-0.01	0.40

Table S24: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was Laplace with negative correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	Mean	Bias	MSE
		$m = 6$		
1	5	-0.156	1.71	6782.50
3	5	-0.098	-36.12	10958.32
0.5/1.5	2/8	-0.128	-16.86	6388.38
1/5	2/8	-0.123	-19.86	10595.53
		$m = 500$		
1	5	-0.066	-56.74	113.16
3	5	-0.098	-36.37	123.96
0.5/1.5	2/8	-0.070	-54.29	107.62
1/5	2/8	-0.094	-38.67	119.46

Uniform

Table S25: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression ($\hat{\tau}_1, \hat{\tau}_2$) and fixed intercept logistic regression ($\hat{\beta}_1, \hat{\beta}_2$) and center-specific AUC ($AUC_c(\cdot)$) across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was uniform with negative correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
$m = 6$																			
1	5	0.102	-10.98	57.99	0.120	4.79	56.31	0.467	-4.25	72.80	0.485	-0.49	68.84	0.651	-0.18	0.43	0.651	-0.16	0.43
3	5	0.106	-7.54	88.24	0.122	6.63	88.64	0.468	-4.10	92.29	0.484	-0.70	89.14	0.651	-0.24	0.74	0.651	-0.22	0.73
0.5/1.5	2/8	0.099	-13.52	68.47	0.116	1.67	66.50	0.475	-2.66	69.57	0.492	0.94	68.10	0.651	-0.21	0.40	0.651	-0.18	0.40
1/5	2/8	0.098	-14.16	74.28	0.113	-1.20	72.59	0.475	-2.52	86.25	0.491	0.58	86.29	0.650	-0.28	0.48	0.651	-0.26	0.47
$m = 500$																			
1	5	-0.069	-160.20	342.56	0.116	1.64	6.52	0.320	-34.43	289.25	0.487	-0.18	7.12	0.641	-1.67	1.74	0.653	0.04	0.38
3	5	-0.029	-125.60	217.41	0.113	-1.45	10.91	0.362	-25.68	166.58	0.487	-0.15	10.08	0.647	-0.77	0.95	0.652	-0.01	0.60
0.5/1.5	2/8	-0.070	-161.43	348.14	0.114	0.05	6.73	0.321	-34.18	285.15	0.488	0.04	7.17	0.641	-1.77	1.95	0.652	-0.05	0.36
1/5	2/8	-0.025	-121.96	204.09	0.115	0.61	9.10	0.364	-25.31	161.59	0.486	-0.35	9.48	0.648	-0.65	0.61	0.653	0.05	0.38

Table S26: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was uniform with negative correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	Mean	Bias	MSE
		$m = 6$		
1	5	-0.189	23.32	5972.91
3	5	-0.123	-19.50	12011.31
0.5/1.5	2/8	-0.117	-23.95	6242.38
1/5	2/8	-0.168	9.60	13186.25
		$m = 500$		
1	5	-0.078	-48.97	104.34
3	5	-0.098	-36.07	147.12
0.5/1.5	2/8	-0.077	-50.00	104.96
1/5	2/8	-0.092	-39.73	157.28

S5.2.4 Confounding (No Correlation)

Normal

Table S27: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression ($\hat{\tau}_1, \hat{\tau}_2$) and fixed intercept logistic regression ($\hat{\beta}_1, \hat{\beta}_2$) and center-specific AUC ($AUC_c(\cdot)$) across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was normal with no correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
$m = 6$																			
1	5	0.094	-17.63	67.96	0.112	-1.92	63.76	0.475	-2.57	71.74	0.493	1.15	70.99	0.651	-0.16	0.42	0.651	-0.13	0.41
3	5	0.107	-6.70	86.53	0.119	3.91	87.89	0.475	-2.62	77.75	0.487	-0.09	77.87	0.651	-0.23	0.88	0.651	-0.22	0.87
0.5/1.5	2/8	0.094	-18.22	59.71	0.112	-1.85	55.70	0.479	-1.84	63.76	0.498	2.06	64.17	0.651	-0.15	0.40	0.652	-0.13	0.39
1/5	2/8	0.102	-10.61	75.08	0.115	0.73	73.98	0.482	-1.08	83.00	0.495	1.57	82.51	0.651	-0.21	0.47	0.651	-0.19	0.46
$m = 500$																			
1	5	-0.037	-132.35	236.82	0.115	0.17	7.93	0.356	-26.92	179.50	0.490	0.50	7.87	0.646	-0.91	0.80	0.652	-0.05	0.38
3	5	0.013	-88.86	112.09	0.112	-1.69	9.06	0.407	-16.60	74.16	0.488	0.12	9.12	0.651	-0.24	0.78	0.653	0.03	0.71
0.5/1.5	2/8	-0.040	-134.79	244.34	0.112	-2.37	7.12	0.354	-27.47	186.58	0.488	0.06	7.63	0.646	-0.97	0.80	0.652	-0.07	0.34
1/5	2/8	0.012	-89.52	114.15	0.115	0.58	9.31	0.402	-17.49	80.98	0.487	-0.16	8.24	0.651	-0.25	0.45	0.653	0.03	0.40

Table S28: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was normal with no correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	Mean	Bias	MSE
		$m = 6$		
1	5	-0.145	-5.77	4014.54
3	5	-0.085	-44.64	8173.21
0.5/1.5	2/8	-0.118	-22.96	4569.70
1/5	2/8	-0.135	-12.10	8494.52
		$m = 500$		
1	5	-0.089	-41.72	80.52
3	5	-0.115	-25.10	106.78
0.5/1.5	2/8	-0.086	-43.93	80.25
1/5	2/8	-0.108	-29.90	93.79

Gumbel

Table S29: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression ($\hat{\tau}_1, \hat{\tau}_2$) and fixed intercept logistic regression ($\hat{\beta}_1, \hat{\beta}_2$) and center-specific AUC ($AUC_c(\cdot)$) across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was Gumbel with no correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
								$m = 6$											
1	5	0.103	-9.97	64.81	0.121	5.68	63.48	0.476	-2.36	70.37	0.494	1.30	68.58	0.650	-0.32	0.49	0.650	-0.30	0.47
3	5	0.099	-13.60	80.38	0.112	-2.54	78.45	0.485	-0.48	80.35	0.498	2.18	82.65	0.651	-0.21	1.32	0.651	-0.19	1.32
0.5/1.5	2/8	0.095	-16.56	61.39	0.113	-1.01	55.83	0.476	-2.39	54.09	0.494	1.30	53.26	0.651	-0.19	0.37	0.651	-0.17	0.36
1/5	2/8	0.101	-11.88	79.96	0.113	-1.37	78.44	0.484	-0.72	77.77	0.496	1.78	80.06	0.651	-0.23	0.44	0.651	-0.21	0.43
								$m = 500$											
1	5	-0.040	-134.61	244.86	0.116	1.75	7.18	0.348	-28.58	200.87	0.487	-0.04	7.11	0.646	-1.00	0.89	0.652	-0.06	0.41
3	5	0.010	-91.31	117.52	0.115	0.93	9.07	0.402	-17.57	81.64	0.490	0.37	8.36	0.650	-0.34	2.04	0.652	-0.04	2.05
0.5/1.5	2/8	-0.043	-137.51	254.84	0.114	-0.64	7.15	0.350	-28.23	197.58	0.490	0.47	7.92	0.646	-1.01	0.94	0.652	-0.02	0.35
1/5	2/8	0.008	-93.29	121.66	0.116	1.63	7.42	0.396	-18.87	92.01	0.486	-0.39	8.45	0.650	-0.37	0.71	0.652	-0.07	0.62

Table S30: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was Gumbel with no correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	Mean	Bias	MSE
		$m = 6$		
1	5	-0.169	10.38	4880.57
3	5	-0.176	14.78	8365.69
0.5/1.5	2/8	-0.094	-38.70	4443.63
1/5	2/8	-0.192	25.14	7828.69
		$m = 500$		
1	5	-0.115	-25.12	45.98
3	5	-0.201	30.94	105.51
0.5/1.5	2/8	-0.113	-26.53	46.49
1/5	2/8	-0.190	24.05	88.67

Laplace

Table S31: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression $(\hat{\tau}_1, \hat{\tau}_2)$ and fixed intercept logistic regression $(\hat{\beta}_1, \hat{\beta}_2)$ and center-specific AUC $(AUC_c(\cdot))$ across 500 simulations when center was a confounder and the joint distribution of $\text{logit}(\gamma_c)$ and $f(c)$ was Laplace with no correlation between $\text{logit}(\gamma_c)$ and $f(c)$. The number of centers in the training data, m, the variance of $\text{logit}(\gamma_c)$, $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
$m = 6$																			
1	5	0.099	-13.57	57.03	0.116	1.77	54.96	0.476	-2.45	71.64	0.493	1.17	70.85	0.651	-0.20	0.43	0.651	-0.19	0.43
3	5	0.103	-10.37	72.34	0.115	0.16	70.58	0.479	-1.69	73.54	0.492	0.80	72.02	0.650	-0.38	1.89	0.650	-0.36	1.87
0.5/1.5	2/8	0.095	-16.63	62.00	0.112	-2.16	57.03	0.476	-2.47	62.29	0.492	0.95	60.65	0.651	-0.20	0.40	0.651	-0.18	0.39
1/5	2/8	0.097	-14.82	78.76	0.109	-4.85	76.88	0.485	-0.62	65.74	0.496	1.70	66.42	0.651	-0.24	0.45	0.651	-0.23	0.44
$m = 500$																			
1	5	-0.042	-136.98	252.97	0.115	0.68	7.31	0.347	-28.75	203.83	0.488	0.05	7.68	0.646	-0.98	1.00	0.653	0.03	0.45
3	5	0.007	-93.68	123.05	0.116	0.97	8.52	0.399	-18.19	87.28	0.488	0.16	8.72	0.650	-0.40	0.71	0.652	-0.08	0.62
0.5/1.5	2/8	-0.046	-140.02	264.06	0.114	-0.78	7.00	0.347	-28.84	204.67	0.489	0.29	7.05	0.645	-1.12	1.01	0.652	-0.07	0.38
1/5	2/8	0.002	-98.19	134.61	0.116	1.39	8.26	0.391	-19.77	101.73	0.486	-0.28	8.69	0.649	-0.45	2.67	0.652	-0.10	4.16

Table S32: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was Laplace with no correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	Mean	Bias	MSE
		$m = 6$		
1	5	-0.149	-2.80	4468.99
3	5	-0.146	-5.05	8138.90
0.5/1.5	2/8	-0.171	11.38	3896.49
1/5	2/8	-0.184	20.28	8136.75
		$m = 500$		
1	5	-0.088	-42.38	76.42
3	5	-0.105	-31.57	96.69
0.5/1.5	2/8	-0.085	-44.29	78.80
1/5	2/8	-0.110	-28.03	83.38

Uniform

Table S33: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression $(\hat{\tau}_1, \hat{\tau}_2)$ and fixed intercept logistic regression $(\hat{\beta}_1, \hat{\beta}_2)$ and center-specific AUC $(AUC_c(\cdot))$ across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was uniform with no correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
$m = 6$																			
1	5	0.097	-15.16	61.43	0.116	1.24	57.48	0.480	-1.57	57.54	0.499	2.31	59.42	0.652	-0.11	0.44	0.652	-0.08	0.43
3	5	0.098	-13.96	84.27	0.111	-2.74	83.95	0.484	-0.70	86.00	0.498	2.05	88.39	0.650	-0.34	0.72	0.650	-0.32	0.71
0.5/1.5	2/8	0.094	-17.86	61.48	0.112	-1.69	58.08	0.477	-2.12	64.58	0.496	1.72	64.04	0.651	-0.20	0.37	0.651	-0.18	0.35
1/5	2/8	0.106	-7.23	73.93	0.119	3.82	74.54	0.473	-3.10	77.24	0.485	-0.51	75.34	0.651	-0.25	0.45	0.651	-0.24	0.45
$m = 500$																			
1	5	-0.034	-129.48	225.53	0.114	-0.04	6.60	0.358	-26.69	174.88	0.489	0.26	6.37	0.647	-0.81	0.71	0.652	-0.03	0.37
3	5	0.017	-84.73	103.26	0.114	-0.41	10.00	0.409	-16.16	71.60	0.488	0.04	10.74	0.651	-0.28	0.67	0.652	-0.03	0.63
0.5/1.5	2/8	-0.034	-129.88	227.15	0.114	-0.27	7.17	0.358	-26.69	176.02	0.489	0.20	6.99	0.647	-0.84	0.74	0.652	-0.03	0.37
1/5	2/8	0.019	-83.18	99.77	0.117	2.16	9.79	0.410	-15.95	69.06	0.489	0.30	9.01	0.651	-0.23	0.40	0.652	-0.00	0.37

Table S34: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was uniform with no correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	Mean	Bias	MSE
		$m = 6$		
1	5	-0.135	-11.94	4635.71
3	5	-0.131	-14.80	8795.47
0.5/1.5	2/8	-0.210	37.05	4809.41
1/5	2/8	-0.139	-9.46	8070.33
		$m = 500$		
1	5	-0.093	-39.53	72.06
3	5	-0.121	-21.22	110.96
0.5/1.5	2/8	-0.098	-36.02	66.91
1/5	2/8	-0.108	-29.45	108.98

S5.2.5 Confounding (Positive Correlation)

Normal

Table S35: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression ($\hat{\tau}_1, \hat{\tau}_2$) and fixed intercept logistic regression ($\hat{\beta}_1, \hat{\beta}_2$) and center-specific AUC ($AUC_c(\cdot)$) across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was normal with positive correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
								$m = 6$											
1	5	0.094	-17.81	68.49	0.116	1.46	63.20	0.473	-2.99	63.87	0.495	1.54	60.99	0.652	-0.12	0.44	0.652	-0.09	0.43
3	5	0.105	-7.92	72.88	0.114	-0.51	75.35	0.487	-0.12	74.90	0.496	1.68	75.67	0.651	-0.15	1.06	0.651	-0.15	1.06
0.5/1.5	2/8	0.094	-18.10	58.61	0.114	-0.16	54.42	0.467	-4.27	66.48	0.487	-0.05	61.71	0.651	-0.22	0.43	0.651	-0.19	0.42
1/5	2/8	0.113	-1.54	73.03	0.120	5.22	73.56	0.484	-0.81	81.52	0.491	0.76	82.16	0.651	-0.19	0.47	0.651	-0.18	0.47
								$m = 500$											
1	5	-0.002	-101.54	141.29	0.114	-0.25	7.16	0.388	-20.53	107.01	0.488	-0.02	7.67	0.650	-0.42	0.58	0.652	-0.04	0.46
3	5	0.067	-41.66	32.03	0.114	-0.71	10.47	0.458	-6.08	17.83	0.487	-0.08	10.50	0.652	0.00	0.84	0.653	0.05	0.85
0.5/1.5	2/8	-0.003	-102.86	144.86	0.114	-0.44	6.84	0.386	-20.76	109.18	0.488	0.04	8.11	0.650	-0.38	0.36	0.652	0.02	0.29
1/5	2/8	0.062	-45.88	35.17	0.116	1.53	8.35	0.452	-7.42	21.45	0.488	-0.03	9.01	0.651	-0.14	0.42	0.652	-0.07	0.41

Table S36: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was normal with positive correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	Mean	Bias	MSE
		$m = 6$		
1	5	-0.146	-4.95	2260.53
3	5	-0.151	-1.30	3965.97
0.5/1.5	2/8	-0.127	-17.01	1996.49
1/5	2/8	-0.118	-23.01	4230.00
		$m = 500$		
1	5	-0.106	-30.95	43.80
3	5	-0.131	-14.67	64.60
0.5/1.5	2/8	-0.100	-34.86	51.94
1/5	2/8	-0.140	-8.82	51.07

Gumbel

Table S37: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression ($\hat{\tau}_1, \hat{\tau}_2$) and fixed intercept logistic regression ($\hat{\beta}_1, \hat{\beta}_2$) and center-specific AUC ($AUC_c(\cdot)$) across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was Gumbel with positive correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
									$m = 6$										
1	5	0.100	-12.72	68.96	0.121	5.86	66.10	0.464	-4.79	70.43	0.485	-0.46	63.72	0.651	-0.22	0.40	0.651	-0.19	0.39
3	5	0.107	-6.85	71.03	0.115	0.92	70.04	0.487	-0.23	81.76	0.496	1.61	82.96	0.651	-0.14	1.30	0.651	-0.13	1.29
0.5/1.5	2/8	0.097	-15.60	66.69	0.118	2.73	62.27	0.469	-3.86	65.67	0.490	0.47	59.93	0.651	-0.26	0.38	0.651	-0.22	0.36
1/5	2/8	0.113	-1.42	69.47	0.120	4.84	68.14	0.489	0.26	66.90	0.496	1.75	66.14	0.651	-0.17	0.50	0.651	-0.16	0.50
									$m = 500$										
1	5	-0.007	-105.85	152.93	0.114	-0.67	7.08	0.383	-21.41	115.65	0.488	0.03	7.83	0.649	-0.44	0.52	0.652	0.02	0.41
3	5	0.059	-48.03	38.18	0.114	-0.35	8.84	0.452	-7.28	21.40	0.489	0.35	9.71	0.652	-0.04	0.64	0.653	0.02	0.64
0.5/1.5	2/8	-0.005	-104.42	149.67	0.118	2.78	7.25	0.381	-21.83	120.33	0.488	0.08	8.02	0.649	-0.49	0.50	0.652	-0.06	0.39
1/5	2/8	0.054	-52.82	44.12	0.114	0.05	7.91	0.446	-8.62	25.28	0.488	0.09	8.15	0.652	-0.12	0.41	0.652	-0.04	0.41

Table S38: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was Gumbel with positive correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	Mean	Bias	MSE
		$m = 6$		
1	5	-0.153	-0.53	2190.06
3	5	-0.153	-0.03	3729.06
0.5/1.5	2/8	-0.145	-5.66	2127.58
1/5	2/8	-0.105	-31.23	4403.84
		$m = 500$		
1	5	-0.126	-18.00	32.51
3	5	-0.215	40.42	87.65
0.5/1.5	2/8	-0.127	-16.90	29.13
1/5	2/8	-0.206	34.08	78.03

Laplace

Table S39: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression ($\hat{\tau}_1, \hat{\tau}_2$) and fixed intercept logistic regression ($\hat{\beta}_1, \hat{\beta}_2$) and center-specific AUC ($AUC_c(\cdot)$) across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was Laplace with positive correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
$m = 6$																			
1	5	0.101	-11.97	57.62	0.122	6.30	54.82	0.470	-3.72	64.48	0.490	0.51	59.37	0.651	-0.21	0.53	0.651	-0.18	0.52
3	5	0.107	-6.77	78.73	0.117	2.00	78.58	0.485	-0.60	77.41	0.495	1.48	76.71	0.652	-0.12	3.56	0.652	-0.11	3.56
0.5/1.5	2/8	0.100	-12.35	64.72	0.120	4.45	59.64	0.468	-4.07	68.76	0.487	-0.15	64.83	0.651	-0.22	0.42	0.651	-0.19	0.41
1/5	2/8	0.113	-0.94	59.58	0.122	6.83	60.79	0.485	-0.59	64.34	0.494	1.20	62.39	0.652	-0.12	0.46	0.652	-0.11	0.46
$m = 500$																			
1	5	-0.007	-106.15	154.08	0.116	1.22	7.40	0.380	-22.10	122.75	0.487	-0.06	7.30	0.650	-0.40	0.63	0.653	0.04	0.52
3	5	0.054	-52.74	44.01	0.114	-0.75	8.41	0.446	-8.46	24.48	0.488	0.10	8.18	0.652	-0.07	1.52	0.652	0.01	1.51
0.5/1.5	2/8	-0.011	-109.75	164.43	0.112	-2.23	6.90	0.382	-21.73	119.69	0.490	0.44	8.11	0.649	-0.53	0.45	0.652	-0.04	0.29
1/5	2/8	0.051	-55.39	48.15	0.117	2.60	8.11	0.439	-9.90	30.38	0.488	0.05	7.92	0.652	-0.08	0.44	0.652	0.01	0.43

Table S40: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was Laplace with positive correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	Mean	Bias	MSE
		$m = 6$		
1	5	-0.161	4.93	2432.85
3	5	-0.173	13.03	4223.45
0.5/1.5	2/8	-0.157	2.14	2328.40
1/5	2/8	-0.142	-7.65	4262.70
		$m = 500$		
1	5	-0.102	-33.75	50.83
3	5	-0.133	-13.52	53.08
0.5/1.5	2/8	-0.101	-34.02	48.38
1/5	2/8	-0.128	-16.25	51.94

Uniform

Table S41: Mean, percent bias, and mean squared error (MSE) of the coefficient estimates from random intercept logistic regression ($\hat{\tau}_1, \hat{\tau}_2$) and fixed intercept logistic regression ($\hat{\beta}_1, \hat{\beta}_2$) and center-specific AUC ($AUC_c(\cdot)$) across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was uniform with positive correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	$\hat{\tau}_1$			$\hat{\beta}_1$			$\hat{\tau}_2$			$\hat{\beta}_2$			$AUC_c(\hat{\boldsymbol{\tau}})$			$AUC_c(\hat{\boldsymbol{\beta}})$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
$m = 6$																			
1	5	0.094	-17.73	64.46	0.117	2.39	60.74	0.469	-3.74	62.06	0.493	1.03	59.05	0.651	-0.18	0.49	0.651	-0.14	0.49
3	5	0.110	-3.88	81.96	0.116	1.51	81.13	0.481	-1.28	75.02	0.488	0.11	80.90	0.651	-0.23	0.77	0.651	-0.23	0.78
0.5/1.5	2/8	0.091	-20.31	68.73	0.115	0.42	62.68	0.471	-3.34	59.35	0.495	1.52	58.37	0.651	-0.19	0.38	0.651	-0.16	0.36
1/5	2/8	0.113	-1.32	70.58	0.119	4.18	70.02	0.486	-0.38	72.73	0.492	0.96	76.58	0.652	-0.13	0.44	0.652	-0.12	0.44
$m = 500$																			
1	5	0.001	-99.48	135.59	0.115	0.87	7.32	0.389	-20.21	103.65	0.488	0.15	7.70	0.650	-0.41	0.51	0.652	-0.03	0.40
3	5	0.069	-40.02	29.85	0.111	-2.81	10.12	0.463	-5.00	14.30	0.489	0.23	10.07	0.652	-0.06	0.64	0.652	-0.03	0.63
0.5/1.5	2/8	0.000	-99.80	136.88	0.115	0.55	7.90	0.389	-20.33	104.40	0.488	0.04	7.62	0.650	-0.38	0.43	0.652	-0.03	0.35
1/5	2/8	0.070	-39.24	27.83	0.116	1.80	8.82	0.459	-5.88	15.84	0.488	0.03	8.65	0.652	-0.11	0.37	0.652	-0.07	0.37

Table S42: Mean, percent bias, and mean squared error (MSE) of the fixed intercept estimate ($\hat{\tau}_0$) from random intercept logistic regression across 500 simulations when center was a confounder and the joint distribution of logit(γ_c) and $f(c)$ was uniform with positive correlation between logit(γ_c) and $f(c)$. The number of centers in the training data, m, the variance of logit(γ_c), $\sigma^2_{\gamma_c}$, and the variance of $f(c)$, $\sigma^2_{f(c)}$, were varied. Some settings included non-constant $\sigma^2_{\gamma_c}$, denoted by "a/b" in the $\sigma^2_{\gamma_c}$ column; in these scenarios, half of the centers had $\sigma^2_{\gamma_c} = a$ and half had $\sigma^2_{\gamma_c} = b$. Likewise, some settings included non-constant $\sigma^2_{f(c)}$. The MSE is multiplied by 10^4.

$\sigma^2_{\gamma_c}$	$\sigma^2_{f(c)}$	Mean	Bias	MSE
		$m = 6$		
1	5	-0.163	6.38	2209.13
3	5	-0.130	-15.12	3874.74
0.5/1.5	2/8	-0.162	5.32	2393.64
1/5	2/8	-0.131	-14.28	4849.74
		$m = 500$		
1	5	-0.107	-30.49	48.02
3	5	-0.140	-8.83	74.08
0.5/1.5	2/8	-0.103	-32.74	50.48
1/5	2/8	-0.133	-13.28	65.31

S6 Simulation Code

S6.1 Ignoring Center

The following R function was used to conduct the simulations reported in Sections 6.1 and S5.1.

```
library(survival)
library(MASS)
library(rms)

expit <- function(x) exp(x)/(1+exp(x))

simcenter <- function(DMassoc1, DMassoc2, trainM, centersz, corrCMCD,
                        CDassocVar, CMassocVar, corr, CMassocTF, CDassocTF,
                        overallprev, testszC, testszM, Clogit){
        mu_x<-sqrt(2)*qnorm(DMassoc1)
        mu_y<-sqrt(2)*qnorm(DMassoc2)

        CDassocVar <- rep(CDassocVar,2)
                        ### x 2 for conditional and marginal AUC test sets
        CMassocVar <- rep(CMassocVar,2)

        ### Create datasets
        # Centers
        center <- rep(1:trainM,times=centersz)
```

```
centerTEC <- rep(1:trainM,times=testszC)
centerTEM <- rep(1:trainM,times=testszM)

# Disease indicators
# assign gamma_c and f(c)
if(CMassocTF & CDassocTF){
        CDCMvals <- mvrnorm(2*trainM, c(0,0),
                matrix(c(1,corrCMCD,corrCMCD,1),nrow=2,byrow=T)) *
                cbind(sqrt(CDassocVar),sqrt(CMassocVar))
        CDassoc <- expit(CDCMvals[,1])
        CMassoc <- CDCMvals[,2]
}else if(CDassocTF){
        CDCMvals <- rnorm(2*trainM, 0, sqrt(CDassocVar))
        CDassoc <- expit(CDCMvals)
        CMassoc <- rep(0, 2*trainM)
}else if(CMassocTF){
        CDCMvals <- rnorm(2*trainM, 0, sqrt(CMassocVar))
        CMassoc <- CDCMvals
        CDassoc <- rep(overallprev, 2*trainM)
}

centerprTR <- rep(CDassoc[1:trainM], times=centersz)
disTR <- rbinom(length(centerprTR),1,centerprTR)
fcTR <- rep(CMassoc[1:trainM], times=centersz)
```

```
centerprTEC <- rep(CDassoc[(trainM+1):(2*trainM)], times=testszC)
disTEC <- rbinom(length(centerprTEC),1,centerprTEC)
fcTEC <- rep(CMassoc[(trainM+1):(2*trainM)], times=testszC)

centerprTEM <- rep(CDassoc[(trainM+1):(2*trainM)], times=testszM)
disTEM <- rbinom(length(centerprTEM),1,centerprTEM)
fcTEM <- rep(CMassoc[(trainM+1):(2*trainM)], times=testszM)

 # Marker values
meanvecTR<-cbind(mu_x*disTR + fcTR, mu_y*disTR + fcTR)
markersTR <- mvrnorm(length(centerprTR),c(0,0),
                matrix(c(1,corr,corr,1),nrow=2,byrow=T)) + meanvecTR
train <- data.frame("dis"=disTR,"x"=markersTR[,1],
                "y"=markersTR[,2],"center"=center)

meanvecTEC<-cbind(mu_x*disTEC + fcTEC, mu_y*disTEC + fcTEC)
markersTEC <- mvrnorm(length(centerprTEC),c(0,0),
                matrix(c(1,corr,corr,1),nrow=2,byrow=T)) + meanvecTEC
testC <- data.frame("dis"=disTEC,"x"=markersTEC[,1],
                "y"=markersTEC[,2],"center"=centerTEC)

meanvecTEM<-cbind(mu_x*disTEM + fcTEM, mu_y*disTEM + fcTEM)
markersTEM <- mvrnorm(length(centerprTEM),c(0,0),
                matrix(c(1,corr,corr,1),nrow=2,byrow=T)) + meanvecTEM
testM <- data.frame("dis"=disTEM,"x"=markersTEM[,1],
```

```
                "y"=markersTEM[,2],"center"=centerTEM)

### Fit regressions to this dataset
train$centerFAC<-as.factor(train$center)
testC$centerFAC<-as.factor(testC$center)
testM$centerFAC<-as.factor(testM$center)

# Ignoring center
igGLM <- glm(dis ~ x + y, data=train, family="binomial")
# Fixed effect
if(Clogit){
        feGLM<-clogit(dis ~ x + y + strata(centerFAC), data=train)
}else{
        feGLM<-glm(dis ~ x + y + centerFAC, data=train,
                        family=binomial)
}

# True parameter values
betaX <- (mu_x-corr*mu_y)/(1-corr^2)
betaY <- (mu_y-corr*mu_x)/(1-corr^2)

### Apply regression results to population and get AUC
# Conditional
predIGC <- data.frame(testC,"predvals"=predict(igGLM,newdata=testC))
if(Clogit){
```

```
        predFEC <- data.frame(testC,
                        "predvals"=apply(testC,1,function(k)
                        as.numeric(k["x"])*feGLM$coef[1]+
                        as.numeric(k["y"])*feGLM$coef[2]))
}else{
        predFEC <- data.frame(testC,
                        "predvals"=apply(testC,1,function(k)
                        as.numeric(k["x"])*feGLM$coef[2]+
                        as.numeric(k["y"])*feGLM$coef[3]))
}

# Marginal
predIGM <- data.frame(testM,"predvals"=predict(igGLM,newdata=testM))
if(Clogit){
        predFEM <- data.frame(testM,
                        "predvals"=apply(testM,1,function(k)
                        as.numeric(k["x"])*feGLM$coef[1]+
                        as.numeric(k["y"])*feGLM$coef[2]))
}else{
        predFEM <- data.frame(testM,
                        "predvals"=apply(testM,1,function(k)
                        as.numeric(k["x"])*feGLM$coef[2]+
                        as.numeric(k["y"])*feGLM$coef[3]))
}
```

```
    # Get aAUC
    # naive AUC based on each set of predicted values
    margIG <- somers2(predIGM$predvals, predIGM$dis)[1]
    margFE <- somers2(predFEM$predvals, predFEM$dis)[1]

    # marginal AUC based on each set of predicted values
    condIG <- somers2(predIGC[predIGC$center==1,]$predvals,
                        predIGC[predIGC$center==1,]$dis)[1]
    condFE <- somers2(predFEC[predFEC$center==1,]$predvals,
                        predFEC[predFEC$center==1,]$dis)[1]

    ## True AUC
    alpha1 <- (mu_x - corr*mu_y)/(1 - corr^2)
    alpha2 <- (mu_y - corr*mu_x)/(1 - corr^2)
    AUCnum <- mu_x + (alpha2/alpha1)*mu_y
    AUCdenom <- sqrt( 2*(1 + ((alpha2^2)/(alpha1^2))
                        + (2*alpha2*corr)/alpha1) )
    trueAUC <- pnorm(AUCnum/AUCdenom)

    c("coefIG"=coef(igGLM), "margIG"=margIG, "condIG"=condIG,
      "coefFE"=coef(feGLM), "margFE"=margFE, "condFE"=condFE,
      "truebetaX"=betaX, "truebetaY"=betaY, "trueAUC"=trueAUC)
}

### Example
```

```
simcenter(DMassoc1=0.6, DMassoc2=0.65, trainM=6, centersz=rep(200,6),
                        corrCMCD=-0.75, CDassocVar=rep(1,6), CMassocVar=rep(5,6),
                        corr=0.5, CMassocTF=TRUE, CDassocTF=TRUE, overallprev=0.5,
                        testszC=c(200000,rep(0,5)), testszM=rep(30000,6),
                        Clogit=FALSE)
## conditional AUC evaluated in 1 test center (because testszC=c(200000,rep(0,5)))
```

S6.2 RILR vs. FILR

The following R function was used to conduct the simulations reported in Sections 6.2 and S5.2.

```
library(survival)
library(lme4)
library(MASS)
library(rms)
library(VGAM)
library(QRM)
library(copula)

expit <- function(x) exp(x)/(1+exp(x))

simcenter <- function(DMassoc1, DMassoc2, trainM, centersz, RIdist,
                        corrCMCD, CDassocVar, CMassocVar, corr, CMassocTF,
                        CDassocTF, overallprev, testsz, Clogit){
        mu_x<-sqrt(2)*qnorm(DMassoc1)
```

```
mu_y<-sqrt(2)*qnorm(DMassoc2)

### Create datasets
# Centers
center <- rep(1:trainM,times=centersz)
centerTE <- rep(trainM+1, testsz) ## test center is the (trainM+1)th
## Disease indicators
# get gamma_c and f(c)
if(CMassocTF & CDassocTF){
        if(RIdist=="normal"){
                CDCMvals <- mvrnorm(trainM+1, c(0,0),
                        matrix(c(1,corrCMCD,corrCMCD,1),nrow=2,byrow=T)) *
                        cbind(sqrt(CDassocVar),sqrt(CMassocVar))
        }else if(RIdist=="laplace"){
                tmp <- normalCopula( corrCMCD, dim=2 )
                x <- rCopula(trainM+1, tmp)
                CDCMvals <- cbind( qlaplace(x[,1], 0, sqrt(CDassocVar/2)),
                        qlaplace(x[,2], 0, sqrt(CMassocVar/2)) )
        }else if(RIdist=="gumbel"){
                tmp <- normalCopula( corrCMCD, dim=2 )
                x <- rCopula(trainM+1, tmp)
                CDCMvals <- cbind(qGumbel(x[,1],
                        digamma(1)*sqrt(6*CDassocVar/(pi^2)),
                                sqrt(6*CDassocVar/(pi^2))),
                        qGumbel(x[,2], digamma(1)*sqrt(6*CMassocVar/(pi^2)),
```

```
                                sqrt(6*CMassocVar/(pi^2))))
        }else if(RIdist=="uniform"){
                tmp <- normalCopula( corrCMCD, dim=2 )
                x <- rCopula(trainM+1, tmp)
                CDCMvals <- cbind(qunif(x[,1], -sqrt(3*CDassocVar),
                                                sqrt(3*CDassocVar)),
                                qunif(x[,2], -sqrt(3*CMassocVar),
                                                sqrt(3*CMassocVar)))
        }
        CDassoc <- expit(CDCMvals[,1])
        CMassoc <- CDCMvals[,2]
}else if(CDassocTF){
        if(RIdist=="normal"){
                CDCMvals <- rnorm(trainM+1, 0, sqrt(CDassocVar))
        }else if(RIdist=="laplace"){
                CDCMvals <- rlaplace(trainM+1, 0, sqrt(CDassocVar/2))
        }else if(RIdist=="gumbel"){
                CDCMvals <- rGumbel(trainM+1,
                                digamma(1)*sqrt(6*CDassocVar/(pi^2)),
                                        sqrt(6*CDassocVar/(pi^2)) )
        }else if(RIdist=="uniform"){
                CDCMvals <- runif(trainM+1, -sqrt(3*CDassocVar),
                                        sqrt(3*CDassocVar))
        }
        CDassoc <- expit(CDCMvals)
```

```
        CMassoc <- rep(0, trainM+1)
}else if(CMassocTF){
        if(RIdist=="normal"){
                CDCMvals <- rnorm(trainM+1, 0, sqrt(CMassocVar))
        }else if(RIdist=="laplace"){
                CDCMvals <- rlaplace(trainM+1, 0, sqrt(CMassocVar/2))
        }else if(RIdist=="gumbel"){
                CDCMvals <- rGumbel(trainM+1,
                                digamma(1)*sqrt(6*CMassocVar/(pi^2)),
                                        sqrt(6*CMassocVar/(pi^2)) )
        }else if(RIdist=="uniform"){
                CDCMvals <- runif(trainM+1, -sqrt(3*CMassocVar),
                                        sqrt(3*CMassocVar))
        }
        CMassoc <- CDCMvals
        CDassoc <- rep(overallprev, trainM+1)
}

centerpr <- rep(CDassoc, times=c(centersz,testsz))
dis <- rbinom(length(centerpr),1,centerpr)
fc <- rep(CMassoc, times=c(centersz,testsz))

 # Marker values
meanvec<-cbind(mu_x*dis + fc, mu_y*dis + fc)
markers <- mvrnorm(length(centerpr),c(0,0),
```

```
                matrix(c(1,corr,corr,1),nrow=2,byrow=T)) + meanvec
popdat <- data.frame("dis"=dis,"x"=markers[,1],"y"=markers[,2],
                "center"=c(center,centerTE))

train <- popdat[which(popdat$center %in% 1:trainM),]
test <- popdat[which(popdat$center==(trainM+1)),]

### Fit regressions to this dataset
train$centerFAC<-as.factor(train$center)
test$centerFAC<-as.factor(test$center)

# Ignoring center
igGLM <- glm(dis ~ x + y, data=train, family="binomial")
# Fixed effect
if(Clogit){
        feGLM<-clogit(dis ~ x + y + strata(centerFAC), data=train)
}else{
        feGLM<-glm(dis ~ x + y + centerFAC, data=train,
                        family=binomial)
}

# Random intercept
REGLM<-glmer(dis ~ x + y + (1 | centerFAC),data = train,
                        family = binomial)
```

```
# True parameter values
betaX <- (mu_x-corr*mu_y)/(1-corr^2)
betaY <- (mu_y-corr*mu_x)/(1-corr^2)

### Apply regression results to population and get aAUC
# Get predicted values
predIG <- data.frame(test,"predvals"=predict(igGLM,newdata=test))
if(Clogit){
        predFE <- data.frame(test,"predvals"=apply(test,1,function(k)
                as.numeric(k["x"])*feGLM$coef[1]+
                as.numeric(k["y"])*feGLM$coef[2]))
}else{
        predFE <- data.frame(test,"predvals"=apply(test,1,function(k)
                as.numeric(k["x"])*feGLM$coef[2]+
                as.numeric(k["y"])*feGLM$coef[3]))
}
predRE <- data.frame(test,"predvals"=predict(REGLM,re.form=NA,
                newdata=test))
# Get aAUC
# naive AUC based on each set of predicted values
naiveIG <- somers2(predIG$predvals, predIG$dis)[1]
naiveFE <- somers2(predFE$predvals, predFE$dis)[1]
naiveRE <- somers2(predRE$predvals, predRE$dis)[1]

## True AUC
```

```
alpha1 <- (mu_x - corr*mu_y)/(1 - corr^2)
alpha2 <- (mu_y - corr*mu_x)/(1 - corr^2)
AUCnum <- mu_x + (alpha2/alpha1)*mu_y
AUCdenom <- sqrt( 2*(1 + ((alpha2^2)/(alpha1^2)) +
                        (2*alpha2*corr)/alpha1) )
trueAUC <- pnorm(AUCnum/AUCdenom)

## Concordant centers in sample
concord_train <- sum( sapply(split(train, train[,"center"]),
                function(x) (sum(x[,"dis"])==length(x[,"dis"]) ||
                sum(x[,"dis"])==0) ) )
concord_test <- sum( sapply(split(test, test[,"center"]),
                function(x) (sum(x[,"dis"])==length(x[,"dis"]) ||
                sum(x[,"dis"])==0) ) )

c("coefIG"=coef(igGLM), "naiveIG"=naiveIG,
  "coefFE"=coef(feGLM), "naiveFE"=naiveFE,
  "coefRE"=REGLM@beta, "naiveRE"=naiveRE,
  "truebetaX"=betaX, "truebetaY"=betaY, "trueAUC"=trueAUC,
  "concord_train"=concord_train, "concord_test"=concord_test,
  "RIdist"=RIdist, "corrCMCD"=corrCMCD, "trainM"=trainM,
  "overallprev"=overallprev,
  "CDassocVar" = ifelse(CDassocTF,
                        ifelse(min(CDassocVar) != max(CDassocVar),
                        paste(min(CDassocVar),"/",max(CDassocVar),sep=""),
```

```
                                        min(CDassocVar)),NA),
          "CMassocVar" = ifelse(CMassocTF,
                                ifelse(min(CMassocVar) != max(CMassocVar),
                                paste(min(CMassocVar),"/",max(CMassocVar),sep=""),
                                        min(CMassocVar)),NA) )
}


### Example
simcenter(DMassoc1=0.6, DMassoc2=0.65, trainM=500, centersz=rep(20,500),
                RIdist="normal", corrCMCD=-0.5, CDassocVar=rep(1,501),
                CMassocVar=rep(5,501), corr=0.5, CMassocTF=TRUE,
                CDassocTF=TRUE, overallprev=0.5, Clogit=TRUE, testsz=10000)
```

References

1. Han AK. Non-parametric analysis of a generalized regression model: the maximum rank correlation estimator. *J Econom* 1987; 35: 303–316.

2. Pepe MS. *The statistical evaluation of medical tests for classification and prediction.* Oxford: Oxford University Press, 2003. pp. 66–92, 130–166.