SUPPLEMENTARY ARTICLE FOR "IMPROVING EFFICIENCY IN BIOMARKER INCREMENTAL VALUE EVALUATION UNDER TWO-PHASE STUDY DESIGNS"

By Yingye Zheng, Marshall Brown Anna Lok and Tianxi Cai*

Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109
E-mail: yzheng@fhcrc.org*;* mdbrown@fredhutch.org
Division of Gastroenterology, University of Michigan Ann Arbor, MI 48109
E-mail: aslok@med.umich.edu
Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115
E-mail: tcai@hsph.harvard.edu

APPENDIX

A. True Weights based IPW (TIPW). Consider a class of TIPW statistics $\widehat{\mathbb{R}}^{\text{TIPW}} = N^{-\frac{1}{2}} \sum_{i=1}^N \widetilde{\omega}_i \mathbf{R}_i$. The form of $\widetilde{\pi}_i$ can be obtained explicitly for both stratified CCH (sCCH) (Gray, 2009; Liu et al., 2012) and NCC (Samuelsen, 1997; Cai and Zheng, 2012) designs with a discrete stratification/matching variable $\mathbf{Z}$ taking S unique values, $\mathbf{z}_1, ..., \mathbf{z}_S$. Specifically, for sCCH design,

$$\widetilde{\pi}_i = \sum_{d=0}^{1} \sum_{s=1}^{S} \widetilde{\pi}_{ds} I(\delta_i = d, \mathbf{Z}_i = \mathbf{z}_s), \quad \text{where} \quad \widetilde{\pi}_{ds} = \frac{n_{ds}}{\sum_{j=1}^N I(\delta_j = d, \mathbf{Z}_j = \mathbf{z}_s)},$$

and n_{ds} is the number of subjects sampled from the set $\{i : \delta_i = d, \mathbf{Z}_i = \mathbf{z}_s\}$, typically specified by design. For a matched NCC (mNCC) design with m controls matched to each case on the matching variable $\mathbf{Z}$, $\widetilde{\pi}_i$ can be calculated as $\widetilde{\pi}_i = \delta_i + \bar{\delta}_i\{1 - \widetilde{G}(\mathbf{W}_i)\}$, with

$$\widetilde{G}(\mathbf{W}_i) = \prod_{j: X_j \leq X_i} \left\{1 - \frac{m\delta_j I(\mathbf{Z}_j = \mathbf{Z}_i)}{\sum_{k=1}^N I(X_k \geq X_j, \mathbf{Z}_k = \mathbf{Z}_i) - 1}\right\}, \tag{A·1}$$

where throughout we let $\bar{\delta}_i = 1 - \delta_i$ for notational ease.

*To whom correspondence should be addressed.

B. Variance Estimation for the Augmented Estimators. Throughout we assume that the censoring time C has a finite support $[0,\tau]$, which is shorter than that of the event time T with $P(T>\tau)>0$. The markers $\mathbf{Y}$ are assumed to be continuous and bounded with $|\mathbf{Y}|\leq \mathcal{Y}_0<\infty$ and the true parameter values $\boldsymbol{\beta}_0$ are assumed to be interior points of a compact parameter space Ω. In addition, we assume the joint density of $\mathbf{Y}$, T and C has continuous derivatives. We also assume the same regularity conditions as in Andersen & Gill (1982). Unless noted otherwise, the sup over time t is taken over $[0,\tau]$ and the sup over $\boldsymbol{\beta}$ is taken over Ω. We use the notation $\approx$ to denote equivalence up to $o_p(1)$.

For simplicity, we assume that no additional matching variables are used; all cases ($\delta=1$) will be selected into the subcohort and a fraction of controls will be selected according to the study design. We next derive the variance form for the augmented statistic of the form $\widehat{\mathbb{R}}^{\text{AIPW}}=N^{-\frac{1}{2}}\sum_{i=1}^{N}\widehat{w}_i\mathbf{R}_i$, where $E(\mathbf{R}_i)=\mathbf{0}$, $\widehat{\omega}_i=V_i/\widehat{\pi}_i(\mathbf{w}_i)$, as specified in Section 3.

To derive the variance form for $\widehat{\mathbb{R}}^{\text{AIPW}}$, we first note that similar arguments as given in the appendix of Cai and Zheng (2012) can be used to show that $\max_i|\widehat{\pi}_0(\mathbf{w}_i)-\widetilde{\pi}_0(X_i)|\to 0$ and

$$\sup_{\mathbf{w}}|n^{-1}\sum_{i=1}^{N}\bar{\delta}_i\widetilde{\omega}_iK_h(\mathbf{W}_i-\mathbf{w})\{\mathbf{R}_i-\boldsymbol{\mu}_{0\mathbf{R}}(\mathbf{w})\}|\to 0$$

in probability, where $\boldsymbol{\mu}_{0\mathbf{R}}(\mathbf{w})=E(\mathbf{R}_i\mid\mathbf{W}_i=\mathbf{w},\delta_i=0)$, $\widetilde{\pi}_{0i}=\widetilde{\pi}_0(X_i)=P(V_i=1\mid\mathscr{D},\delta_i=0)=P(V_i=1\mid X_i,\delta_i=0,\mathscr{F})$ and $\mathscr{F}=\{(X_i,\delta_i),i=1,...,N\}$. It follows that

$$\begin{aligned}
\widehat{\mathbb{R}}^{\text{AIPW}} &\approx N^{-\frac{1}{2}}\sum_{i=1}^{N}\widetilde{\omega}_i\mathbf{R}_i+N^{-\frac{1}{2}}\sum_{i=1}^{N}\bar{\delta}_i\widetilde{\omega}_i\frac{\widetilde{\pi}_0(X_i)-\widehat{\pi}_0(\mathbf{W}_i)}{\widetilde{\pi}_0(X_i)}\\
&\approx N^{-\frac{1}{2}}\sum_{i=1}^{N}\widetilde{\omega}_i\mathbf{R}_i-N^{-\frac{1}{2}}\sum_{j=1}^{n}(\widetilde{w}_j-1)\left\{N^{-1}\sum_{i=1}^{N}\bar{\delta}_i\widetilde{\omega}_i\mathbf{R}_i\frac{K_h(\mathbf{W}_j-\mathbf{W}_i)}{f_0(\mathbf{W}_j)}\right\}\\
&\approx N^{-\frac{1}{2}}\sum_{i=1}^{N}\widetilde{\omega}_i\mathbf{R}_i-N^{-\frac{1}{2}}\sum_{j=1}^{n}(\widetilde{w}_j-1)\boldsymbol{\mu}_{0\mathbf{R}}(\mathbf{W}_j)
\end{aligned}$$

It follows that

$$\widehat{\mathbb{R}}^{\mathrm{AIPW}} \approx N^{-\frac{1}{2}} \sum_{i=1}^{N} \mathbf{R}_i + N^{-\frac{1}{2}} \sum_{i=1}^{N} (\widetilde{w}_i - 1) \mathbf{R}_{0i}^{\perp}$$

$$= N^{-\frac{1}{2}} \sum_{i=1}^{N} \mathbf{R}_i + N^{-\frac{1}{2}} \sum_{i=1}^{N} \bar{\delta}_i \left(\frac{V_i}{\widetilde{\pi}_{0i}} - 1 \right) \mathbf{R}_{0i}^{\perp}$$

where $\mathbf{R}_{0i}^{\perp} = \mathbf{R}_i - \mu_{0\mathbf{R}}(\mathbf{W}_i)$. Therefore, the variance of $\widehat{\mathbb{R}}$ is asymptotically

$$\boldsymbol{\Sigma}_{\mathbf{R}}^{\mathrm{CCH}} = E(\mathbf{R}_i^{\otimes 2}) + E\left\{ \frac{1 - \pi_0^{\mathrm{CCH}}}{\pi_0^{\mathrm{CCH}}} \bar{\delta}_i \mathrm{Var}(\mathbf{R}_i \mid \mathbf{W}_i^{\mathrm{CCH}}) \right\} \tag{A·2}$$

$$\boldsymbol{\Sigma}_{\mathbf{R}}^{\mathrm{NCC}} = E(\mathbf{R}_i^{\otimes 2}) + E\left\{ \frac{1 - \pi_{0i}^{\mathrm{NCC}}}{\pi_{0i}^{\mathrm{NCC}}} \bar{\delta}_i \mathrm{Var}(\mathbf{R}_i \mid \mathbf{W}_i^{\mathrm{NCC}}) \right\} \tag{A·3}$$

where $\pi_0^{\mathrm{CCH}} = n_0/N_0$, $\pi_{0i}^{\mathrm{NCC}} = 1 - e^{-m\Lambda_{\mathrm{marg}}(X_i)}$ and $\Lambda_{\mathrm{marg}}(u) = \int_0^u \frac{dE\{I(X_i \le v)\delta_i\}}{P(X > v)}$. It is interesting to note that in order to derive the form of $\boldsymbol{\Sigma}_{\mathbf{R}}^{\mathrm{NCC}}$, we used the fact that

$$E\left\{ \mathbf{R}_{0i}^{\perp} I(X_i > u) \bar{\delta}_i \frac{1 - \pi_{0i}^{\mathrm{NCC}}}{\pi_{0i}^{\mathrm{NCC}}} \right\} = E\left[\bar{\delta}_i \frac{1 - \pi_{0i}^{\mathrm{NCC}}}{\pi_{0i}^{\mathrm{NCC}}} E\{\mathbf{R}_{0i}^{\perp} \mid X_i, \delta_i = 0\} \right] = 0$$

Compared to the variance of the corresponding IPW estimators without augmentation, the gain in efficiency due to augmentation can be expressed as

$$\boldsymbol{\Delta}_{\mathbf{R}}^{\mathrm{CCH}} = E\left\{ \frac{1 - \pi_0^{\mathrm{CCH}}}{\pi_0^{\mathrm{CCH}}} \bar{\delta}_i \boldsymbol{\mu}_{0\mathbf{R}}(\mathbf{W}_i^{\mathrm{CCH}})^{\otimes 2} \right\} \ge 0 \tag{A·4}$$

for the CCH design and

$$\boldsymbol{\Delta}_{\mathbf{R}}^{\mathrm{NCC}} = E\left[\frac{1 - \pi_{0i}^{\mathrm{NCC}}}{\pi_{0i}^{\mathrm{NCC}}} \bar{\delta}_i \left\{ \mathbf{R}_i^{\otimes 2} - \int \frac{\eta_{0\mathbf{R}}(u)^{\otimes 2}}{\pi(u)} d\Lambda_{\mathrm{marg}}(u) - (\mathbf{R}_{0i}^{\perp})^{\otimes 2} \right\} \right] \tag{A·5}$$

$$= E\left[\frac{1 - \pi_{0i}^{\mathrm{NCC}}}{\pi_{0i}^{\mathrm{NCC}}} \bar{\delta}_i \left\{ \boldsymbol{\mu}_{0\mathbf{R}}(\mathbf{W}_i^{\mathrm{NCC}})^{\otimes 2} - \int \frac{\eta_{0\mathbf{R}}(u)^{\otimes 2}}{\pi(u)} d\Lambda_{\mathrm{marg}}(u) \right\} \right] \tag{A·6}$$

$$= \lim_{n \to \infty} \mathrm{Var}\left\{ N^{-\frac{1}{2}} \sum_{j=1}^{n} (\widetilde{w}_j - 1) \boldsymbol{\mu}_{0\mathbf{R}}(\mathbf{W}_i^{\mathrm{NCC}}) \right\} \ge 0 \tag{A·7}$$

for the NCC design, where

$$\eta_{0\mathbf{R}}(u) = E\left\{\mathbf{R}_i I(X_i > u)\bar{\delta}_i \frac{1-\pi_{0i}^{\rm NCC}}{\pi_{0i}^{\rm NCC}}\right\} = E\left\{\boldsymbol{\mu}_{0\mathbf{R}}(\mathbf{W}_i^{\rm NCC}) I(X_i > u)\bar{\delta}_i \frac{1-\pi_{0i}^{\rm NCC}}{\pi_{0i}^{\rm NCC}}\right\}.$$

For both settings, the larger variation $\boldsymbol{\mu}_{0\mathbf{R}}(\mathbf{W}_i)$ has, the larger efficiency gain the augmentation achieves.

C. Expansions for AIPW Estimators of Accuracy Parameters. We consider an AIPW profile likelihood estimator of $\boldsymbol{\beta}_0$ as $\widehat{\boldsymbol{\beta}} = \operatorname{argmax}_{\boldsymbol{\beta}} \widehat{\ell}_{\rm profile}(\boldsymbol{\beta})$ as specified in Section 3. It follows from the profile likelihood theory (Murphy et al., 1997) and similar arguments given in Zeng and Lin (2006) that $\widehat{\boldsymbol{\beta}}$ is consistent for $\boldsymbol{\beta}_0$ and the estimated baseline function $\widehat{H}(t) = \widehat{H}(t;\widehat{\boldsymbol{\beta}}) = \sum_{k=1}^{n_1} I(t_k \le t)\Delta\widehat{H}(t_k;\widehat{\boldsymbol{\beta}})$ is uniformly consistent for $H(t)$. Furthermore,

$$\widehat{\mathbb{R}}_{\theta_t} = N^{\frac{1}{2}}(\widehat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_{0t}) = N^{-\frac{1}{2}}\sum_{i=1}^{N}\widehat{\omega}_i \mathbf{R}_{\theta_t i} + o_p(1),$$

where $\widehat{\boldsymbol{\theta}}_t = (\log \widehat{H}(t), \widehat{\boldsymbol{\beta}}^{\mathsf{T}})^{\mathsf{T}}$, $\boldsymbol{\theta}_{t0} = (\log H(t), \boldsymbol{\beta}_0)^{\mathsf{T}}$, and $\mathbf{R}_{\theta_t}$ is the influence functions for $\widehat{\mathbb{R}}_{\theta_t}$ under the full cohort. Similarly, under the reduced model $\mathcal{R}_t^{\rm old}(\mathbf{Y}) = \mathcal{T}_0[\log\{H_{\rm old}(t)\} + \mathbf{Y}_{\rm old}^{\mathsf{T}}\boldsymbol{\alpha}_{\rm old0}]$, let $\widehat{H}_{\rm old}(t)$ and $\widehat{\boldsymbol{\alpha}}_{\rm old}$ denote the resulting respective estimators of $H_{\rm old}(t)$ and $\boldsymbol{\alpha}_{\rm old0}$. Let $\widehat{\boldsymbol{\theta}}_t^{\rm old} = [\log\{\widehat{H}_{\rm old}(t)\}, \widehat{\boldsymbol{\alpha}}_{\rm old}^{\mathsf{T}}, \mathbf{0}_{q\times 1}^{\mathsf{T}}]^{\mathsf{T}}$ and $\boldsymbol{\theta}_{t0}^{\rm old} = [\log\{H_{\rm old}(t)\}, \boldsymbol{\alpha}_{\rm old0}^{\mathsf{T}}, \mathbf{0}_{q\times 1}^{\mathsf{T}}]^{\mathsf{T}}$, where q is the dimension of $\mathbf{Y}_{\rm new}$. Then we may obtain $\widehat{\mathbb{R}}_{\theta_t}^{\rm old} \equiv N^{\frac{1}{2}}(\widehat{\boldsymbol{\theta}}_t^{\rm old} - \boldsymbol{\theta}_{t0}^{\rm old}) \simeq N^{-\frac{1}{2}}\sum_{i=1}^{N}\mathbf{R}_{\theta_t i}^{\rm old}$ for some influence function $\mathbf{R}_{\theta_t}^{\rm old}$, where the last q component of $\mathbf{R}_{\theta_t}^{\rm old}$ is 0.

To derive the distributions for the estimated TPR and FPR functions for the full and reduced model, we first define

$$\mathrm{TPR}_t(c;\boldsymbol{\theta},\boldsymbol{\theta}^*)=\frac{E\left[\bar{\mathcal{T}}_0(\boldsymbol{\theta}^{\mathsf{T}}\vec{\mathbf{Y}}_i)I\{\bar{\mathcal{T}}_0(\vec{\mathbf{Y}}_i^{\mathsf{T}}\boldsymbol{\theta}^*)\geq c\}\right]}{E\{\bar{\mathcal{T}}_0(\boldsymbol{\theta}^{\mathsf{T}}\mathbf{Y}_i)\}},$$

$$\widehat{\mathrm{TPR}}_t(c;\boldsymbol{\theta},\boldsymbol{\theta}^*)=\frac{\sum_{i=1}^N\widehat{w}_i\bar{\mathcal{T}}_0(\boldsymbol{\theta}^{\mathsf{T}}\vec{\mathbf{Y}}_i)I\{\bar{\mathcal{T}}_0(\vec{\mathbf{Y}}_i^{\mathsf{T}}\boldsymbol{\theta}^*)\geq c\}}{\sum_{i=1}^N\widehat{w}_i\bar{\mathcal{T}}_0(\boldsymbol{\theta}^{\mathsf{T}}\vec{\mathbf{Y}}_i)},$$

$$\mathrm{FPR}_t(c;\boldsymbol{\theta},\boldsymbol{\theta}^*)=\frac{E\left[\mathcal{T}_0(\boldsymbol{\theta}^{\mathsf{T}}\mathbf{Y}_i)I\{\bar{\mathcal{T}}_0(\vec{\mathbf{Y}}_i^{\mathsf{T}}\boldsymbol{\theta}^*)\geq c\}\right]}{E\{\mathcal{T}_0(\boldsymbol{\theta}^{\mathsf{T}}\mathbf{Y}_i)\}},$$

$$\widehat{\mathrm{FPR}}_t(c;\boldsymbol{\theta},\boldsymbol{\theta}^*)=\frac{\sum_{i=1}^N\widehat{w}_i\mathcal{T}_0(\boldsymbol{\theta}^{\mathsf{T}}\vec{\mathbf{Y}}_i)I\{\bar{\mathcal{T}}_0(\vec{\mathbf{Y}}_i^{\mathsf{T}}\boldsymbol{\theta}^*)\geq c\}}{\sum_{i=1}^N\widehat{w}_i\mathcal{T}_0(\boldsymbol{\theta}^{\mathsf{T}}\vec{\mathbf{Y}}_i)},$$

where $\bar{\mathcal{T}}_0(x)=1-\mathcal{T}_0(x)$, and for any vector $\boldsymbol{a}$, $\vec{\boldsymbol{a}}=(1,\boldsymbol{a}^{\mathsf{T}})^{\mathsf{T}}$. Then, it is straightforward to see that $\widehat{\mathrm{TPR}}_t^{\mathrm{upd}}(c)=\widehat{\mathrm{TPR}}_t(c;\widehat{\boldsymbol{\theta}}_t,\widehat{\boldsymbol{\theta}}_t)$, $\widehat{\mathrm{FPR}}_t^{\mathrm{upd}}(c)=\widehat{\mathrm{FPR}}_t(c;\widehat{\boldsymbol{\theta}}_t;\widehat{\boldsymbol{\theta}}_t)$, $\mathrm{TPR}_t^{\mathrm{upd}}(c)=\mathrm{TPR}_t(c;\boldsymbol{\theta}_{0t},\boldsymbol{\theta}_{0t})$, and $\mathrm{FPR}_t^{\mathrm{upd}}(c)=\mathrm{FPR}_t(c;\boldsymbol{\theta}_{0t},\boldsymbol{\theta}_{0t})$. On the other hand, $\widehat{\mathrm{TPR}}_t^{\mathrm{old}}(c)=\widehat{\mathrm{TPR}}_t(c;\widehat{\boldsymbol{\theta}}_t,\widehat{\boldsymbol{\theta}}_t^{\mathrm{old}})$, $\widehat{\mathrm{FPR}}_t^{\mathrm{old}}(c)=\widehat{\mathrm{FPR}}_t(c;\widehat{\boldsymbol{\theta}}_t;\widehat{\boldsymbol{\theta}}_t^{\mathrm{old}})$, $\mathrm{TPR}_t^{\mathrm{old}}(c)=\mathrm{TPR}_t(c;\boldsymbol{\theta}_{0t},\boldsymbol{\theta}_{0t}^{\mathrm{old}})$, $\mathrm{FPR}_t^{\mathrm{old}}(c)=\mathrm{FPR}_t(c;\boldsymbol{\theta}_{0t},\boldsymbol{\theta}_{0t}^{\mathrm{old}})$.

It follows from Taylor series expansions and stochastic equicontinuity properties that $\widehat{\mathbb{R}}_{\mathrm{TPR}_t}^{\mathrm{upd}}(c)\equiv N^{\frac{1}{2}}\{\widehat{\mathrm{TPR}}_t^{\mathrm{upd}}(c)-\mathrm{TPR}_t^{\mathrm{upd}}(c)\}$ is asymptotically equivalent to

$$N^{\frac{1}{2}}\{\widehat{\mathrm{TPR}}_t(c;\widehat{\boldsymbol{\theta}}_t,\widehat{\boldsymbol{\theta}}_t)-\mathrm{TPR}_t(c;\widehat{\boldsymbol{\theta}}_t,\widehat{\boldsymbol{\theta}}_t)\}+$$

$$N^{\frac{1}{2}}\{\dot{\mathrm{TPR}}_{1t}(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}_{t0})(\widehat{\boldsymbol{\theta}}_t-\boldsymbol{\theta}_{t0})+\dot{\mathrm{TPR}}_{2t}(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}_{t0})(\widehat{\boldsymbol{\theta}}_t-\boldsymbol{\theta}_{t0})\}$$

$$\simeq N^{\frac{1}{2}}\{\widehat{\mathrm{TPR}}_t(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}_{t0})-\mathrm{TPR}_t^{\mathrm{upd}}(c)+\dot{\mathrm{TPR}}_t(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}_{t0})(\widehat{\boldsymbol{\theta}}_t-\boldsymbol{\theta}_{t0})\}$$

$$\simeq N^{-\frac{1}{2}}\sum_{i=1}^N\widehat{\omega}_i\mathbf{R}_{\mathrm{TPR}_t i}^{\mathrm{upd}}(c),$$

where $\dot{\mathrm{TPR}}_t(c,\boldsymbol{\theta}_1,\boldsymbol{\theta}_2)=\sum_{k=1}^2\dot{\mathrm{TPR}}_{kt}(c;\boldsymbol{\theta}_1,\boldsymbol{\theta}_2)$, $\dot{\mathrm{TPR}}_{kt}(c;\boldsymbol{\theta}_1,\boldsymbol{\theta}_2)=\partial\mathrm{TPR}_t(c;\boldsymbol{\theta}_1,\boldsymbol{\theta}_2)/\partial\boldsymbol{\theta}_k$, and

$$\mathbf{R}^{\text{upd}}_{\text{TPR}_t i}(c) = \dot{\text{TPR}}_t(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}_{t0})^\mathsf{T}\mathbf{R}_{\theta_t i} + \frac{\bar{\mathcal{T}}_0(\boldsymbol{\theta}_{t0}^\mathsf{T}\mathbf{Y}_i)\left[I\left\{\bar{\mathcal{T}}_0(\boldsymbol{\theta}_{t0}^\mathsf{T}\vec{\mathbf{Y}}_i)\geq c\right\} - \text{TPR}^{\text{upd}}_t(c)\right]}{P(T\leq t)}.$$

Similarly, $\widehat{\mathbb{R}}^{\text{upd}}_{\text{FPR}_t}(c) \equiv N^{\frac{1}{2}}\{\widehat{\text{FPR}}^{\text{upd}}_t(c) - \text{FPR}^{\text{upd}}_t(c)\} \simeq N^{-\frac{1}{2}}\sum_{i=1}^N \widehat{\omega}_i \mathbf{R}^{\text{upd}}_{\text{FPR}_t i}(c)$, where

$$R^{\text{upd}}_{\text{FPR}_t i}(c) = \dot{\text{FPR}}_t(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}_{t0})^\mathsf{T}\mathbf{R}_{\theta_t i} + \frac{\mathcal{T}_0(\boldsymbol{\theta}_{t0}^\mathsf{T}\vec{\mathbf{Y}}_i)\left[I\left\{\bar{\mathcal{T}}_0(\boldsymbol{\theta}_{t0}^\mathsf{T}\vec{\mathbf{Y}}_i)\geq c\right\} - \text{FPR}^{\text{upd}}_t(c)\right]}{P(T > t)},$$

$\dot{\text{FPR}}_{kt}(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}_{t0}) = \partial\text{FPR}_t(c;\boldsymbol{\theta}_1,\boldsymbol{\theta}_2)/\partial\boldsymbol{\theta}_k$ and $\dot{\text{FPR}}_t(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}_{t0}) = \sum_{k=1}^2 \dot{\text{FPR}}_{kt}(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}_{t0})$. Following similar lines of argument, we may expand

$$\widehat{\mathbb{R}}^{\text{old}}_{\text{TPR}_t}(c) \equiv N^{\frac{1}{2}}\{\widehat{\text{TPR}}^{\text{old}}_t(c) - \text{TPR}^{\text{old}}_t(c)\} \simeq N^{\frac{1}{2}}\{\widehat{\text{TPR}}_t(c;\widehat{\boldsymbol{\theta}}_t,\widehat{\boldsymbol{\theta}}^{\text{old}}_t) - \text{TPR}_t(c;\widehat{\boldsymbol{\theta}}_t,\widehat{\boldsymbol{\theta}}^{\text{old}}_t)\} +$$

$$N^{\frac{1}{2}}\{\dot{\text{TPR}}_{1t}(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}^{\text{old}}_{t0})(\widehat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_{t0}) + \dot{\text{TPR}}_{2t}(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}^{\text{old}}_{t0})(\widehat{\boldsymbol{\theta}}^{\text{old}}_t - \boldsymbol{\theta}^{\text{old}}_{t0})\}$$

$$\simeq N^{-\frac{1}{2}}\sum_{i=1}^N \mathbf{R}^{\text{old},1}_{\text{TPR}_t i}(c) + N^{-\frac{1}{2}}\sum_{i=1}^N \widehat{\omega}_i\mathbf{R}^{\text{old},2}_{\text{TPR}_t i}(c)$$

and $\widehat{\mathbb{R}}^{\text{old}}_{\text{FPR}_t}(c) \equiv N^{\frac{1}{2}}\{\widehat{\text{FPR}}^{\text{old}}_t(c) - \text{FPR}^{\text{old}}_t(c)\} \simeq N^{-\frac{1}{2}}\sum_{i=1}^N \mathbf{R}^{\text{old},1}_{\text{FPR}_t i}(c) + N^{-\frac{1}{2}}\sum_{i=1}^N \widehat{\omega}_i\mathbf{R}^{\text{old},2}_{\text{FPR}_t i}(c)$, where

$$R^{\text{old},1}_{\text{TPR}_t i}(c) = \dot{\text{TPR}}_{2t}(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}^{\text{old}}_{t0})^\mathsf{T}\mathbf{R}^{\text{old}}_{\theta_t i},$$

$$R^{\text{old},2}_{\text{TPR}_t i}(c) = \dot{\text{TPR}}_{1t}(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}^{\text{old}}_{t0})^\mathsf{T}\mathbf{R}_{\theta_t i} + \frac{\bar{\mathcal{T}}_0(\boldsymbol{\theta}_{t0}^\mathsf{T}\vec{\mathbf{Y}}_i)\left[I\left\{\bar{\mathcal{T}}_0(\vec{\mathbf{Y}}_i^\mathsf{T}\boldsymbol{\theta}^{\text{old}}_{t0})\geq c\right\} - \text{TPR}^{\text{old}}_t(c)\right]}{P(T\leq t)}.$$

$$R^{\text{old},1}_{\text{FPR}_t i}(c) = \dot{\text{FPR}}_{2t}(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}^{\text{old}}_{t0})^\mathsf{T}\mathbf{R}^{\text{old}}_{\theta_t i},$$

$$R^{\text{old},2}_{\text{FPR}_t i}(c) = \dot{\text{FPR}}_{1t}(c;\boldsymbol{\theta}_{t0},\boldsymbol{\theta}^{\text{old}}_{t0})^\mathsf{T}\mathbf{R}_{\theta_t i} + \frac{\mathcal{T}_0(\boldsymbol{\theta}_{t0}^\mathsf{T}\vec{\mathbf{Y}}_i)\left[I\left\{\bar{\mathcal{T}}_0(\vec{\mathbf{Y}}_i^\mathsf{T}\boldsymbol{\theta}^{\text{old}}_{t0})\geq c\right\} - \text{FPR}^{\text{old}}_t(c)\right]}{P(T > t)}.$$

To construct confidence intervals for the accuracy and IncV parameters, we need to estimate the variances of $\widehat{\mathbb{R}}^{\text{upd}}_{\mathcal{A}_t} = N^{\frac{1}{2}}(\widehat{\mathcal{A}}^{\text{upd}}_t - \mathcal{A}^{\text{upd}}_t)$, and $\widehat{\mathbb{R}}^{\text{IncV}}_{\mathcal{A}_t} = N^{\frac{1}{2}}(\widehat{\text{IncV}}_{\mathcal{A}_t} - \text{IncV}_{\mathcal{A}_t})$.

Similarly, we may obtain $\widehat{\mathbb{R}}^{\text{old}}_{\mathcal{A}_t} = N^{\frac{1}{2}}(\widehat{\mathcal{A}}^{\text{old}}_t - \mathcal{A}^{\text{old}}_t) \simeq N^{-\frac{1}{2}}\sum_{i=1}^N \mathbf{R}^{\text{old},1}_{\mathcal{A}_t i} + N^{-\frac{1}{2}}\sum_{i=1}^N \widehat{w}_i \mathbf{R}^{\text{old},2}_{\mathcal{A}_t i}$ combining arguments previously used for such estimators under full cohort studies (Uno et al., 2007; Zheng et al., 2008) and those given here. It then follows from arguments as given in Appendix B that the asymptotic variances of $\widehat{\mathbb{R}}_{\mathcal{A}_t}^{\mathrm{IncV}}$ for the CCH and NCC design are respectively

$$E\left\{(\mathbf{R}_{\mathcal{A}_t i}^{\mathrm{old},1} + \mathbf{R}_{\mathcal{A}_t i}^{\mathrm{old},2} - \mathbf{R}_{\mathcal{A}_t i}^{\mathrm{upd}})^{\otimes 2}\right\} + E\left\{\frac{1-\pi_0^{\mathrm{CCH}}}{\pi_0^{\mathrm{CCH}}}\bar{\delta}_i \mathrm{Var}(\mathbf{R}_{\mathcal{A}_t i}^{\mathrm{old},2} - \mathbf{R}_{\mathcal{A}_t i}^{\mathrm{upd}} \mid \mathbf{W}_i^{\mathrm{CCH}})\right\}$$

$$E\left\{(\mathbf{R}_{\mathcal{A}_t i}^{\mathrm{old},1} + \mathbf{R}_{\mathcal{A}_t i}^{\mathrm{old},2} - \mathbf{R}_{\mathcal{A}_t i}^{\mathrm{upd}})^{\otimes 2}\right\} + E\left\{\frac{1-\pi_{0i}^{\mathrm{NCC}}}{\pi_{0i}^{\mathrm{NCC}}}\bar{\delta}_i \mathrm{Var}(\mathbf{R}_{\mathcal{A}_t i}^{\mathrm{old},2} - \mathbf{R}_{\mathcal{A}_t i}^{\mathrm{upd}} \mid \mathbf{W}_i^{\mathrm{NCC}})\right\}$$

REFERENCES

Cai, T. and Zheng, Y. (2012). Evaluating prognostic accuracy of biomarkers under nested case-control studies. *Biostatistics* **13,** 89–100.

Gray, R. J. (2009). Weighted analyses for cohort sampling designs. *Lifetime Data Analysis* **15,** 24–40.

Liu, D., Cai, T., and Zheng, Y. (2012). Evaluating the predictive value of biomarkers with stratified case-cohort design. *Biometrics* **68,** 1219–1227.

Murphy, S., Rossini, A., and Van der Vaart, A. (1997). Maximum likelihood estimation in the proportional odds model. *Journal of the American Statistical Association* **92,** 968–976.

Samuelsen, S. O. (1997). A pseudolikelihood approach to analysis of nested case-control studies. *Biometrika* **84,** 379–394.

Uno, H., Cai, T., Tian, L., and Wei, L. (2007). Evaluating prediction rules for t-year survivors with censored regression models. *Journal of the American Statistical Association* **102,** 527–537.

Zeng, D. and Lin, D. (2006). Maximum likelihood estimation in semiparametric transformation models for counting processes. *Biometrika* **93,** 627–40.

Zheng, Y., Cai, T., Pepe, M. S., and Levy, W. C. (2008). Time-dependent predictive values of prognostic biomarkers with failure time outcome. *Journal of the American Statistical Association* **103,** 362–368.

Table 1

Results of 5,000 simulations under different CCH and NCC studies. Shown below are bias, empirical standard error (ESE) of the proposed estimators, average of the estimated standard errors (ASE) and coverage percentage (CP).

(a) CCH

		TIPW				AIPW			
	True	% bias	ESE	ASE	CP (%)	% bias	ESE	ASE	CP (%)
AUC	0.88	0.1	0.01	0.01	93.7	0.2	0.01	0.01	91.7
DMR	0.36	0.6	0.03	0.03	95.2	0.7	0.02	0.02	93.0
$TPR(p = 0.05)$	0.95	0.1	0.01	0.01	95.1	0.0	0.01	0.01	95.0
$FPR(p = 0.05)$	0.52	-0.9	0.04	0.05	95.8	-1.2	0.03	0.03	95.2
$PPV(p = 0.05)$	0.27	1.0	0.02	0.02	95.5	0.8	0.01	0.01	94.1
$NPV(p = 0.05)$	0.98	0.0	0.002	0.002	94.5	0.1	0.001	0.002	96.4
$TPR(p = 0.30)$	0.62	-0.1	0.03	0.04	95.6	-0.1	0.03	0.03	94.1
$FPR(p = 0.30)$	0.09	-1.6	0.01	0.01	95.5	-1.9	0.01	0.01	95.3
$PPV(p = 0.30)$	0.57	0.8	0.03	0.03	95.7	0.8	0.02	0.02	95.3
$NPV(p = 0.30)$	0.92	0.0	0.01	0.01	96.1	0.0	0.01	0.01	95.8
$PCF(0.20)$	0.66	0.1	0.03	0.03	95.1	0.3	0.02	0.02	93.6
$PNF(0.85)$	0.62	0.2	0.03	0.03	94.7	0.3	0.02	0.02	93.4

(b) NCC

AUC	0.88	-0.2	0.01	0.01	92.9	-0.2	0.01	0.01	93.4
DMR	0.17	-0.0	0.03	0.04	97.6	-1.1	0.03	0.03	97.7
$TPR(p = 0.01)$	0.93	-0.7	0.01	0.01	96.8	-0.8	0.01	0.01	95.1
$FPR(p = 0.01)$	0.42	-1.6	0.04	0.05	96.9	-2.6	0.04	0.04	96.4
$PPV(p = 0.01)$	0.06	4.1	0.01	0.01	94.9	3.7	0.01	0.01	95.7
$NPV(p = 0.01)$	0.99	-0.0	0.00	0.00	89.7	-0.0	0.000	0.002	100.0
$TPR(p = 0.03)$	0.79	-0.5	0.03	0.03	95.9	-0.6	0.03	0.03	95.7
$FPR(p = 0.03)$	0.19	-1.1	0.02	0.03	95.9	-2.1	0.02	0.02	95.2
$PPV(p = 0.03)$	0.10	3.7	0.01	0.01	95.3	3.2	0.01	0.01	95.4
$NPV(p = 0.03)$	0.99	-0.0	0.001	0.001	94.4	-0.0	0.001	0.002	100.0
$PCF(0.20)$	0.78	-0.3	0.03	0.03	93.8	-0.1	0.03	0.03	93.5
$PNF(0.85)$	0.72	-0.1	0.03	0.03	93.4	0.1	0.03	0.03	93.0

TABLE 2
Comparison of means (empirical SD) for type A: Random CCH designs with different nearest neighbor (NN) tuning parameter values used for the AIPW estimator. Results based on 5,000 simulations.

Setting:	A: Random		
Matching:	none		
Method:	AIPW(δ, Y_1)	AIPW(δ, Y_1)	AIPW(δ, Y_1)
NN Tuning par	0.5	0.7 (default)	0.9
β_1	1.114	1.113	1.112
	(0.0173)	(0.0171)	(0.0171)
β_2	0.697	0.697	0.696
	(0.0228)	(0.0226)	(0.0225)
AUC	0.877	0.876	0.876
	(0.0001)	(0.0001)	(0.0001)
DMR	0.359	0.359	0.359
	(0.0006)	(0.0006)	(0.0006)
$TPR(p = 0.05)$	0.948	0.948	0.948
	(0.0000)	(0.0000)	(0.0000)
$FPR(p = 0.05)$	0.509	0.509	0.510
	(0.0011)	(0.0011)	(0.0011)
$NB(p = 0.05)$	0.135	0.135	0.136
	(0.0001)	(0.0001)	(0.0001)
$PPV(p = 0.05)$	0.271	0.271	0.271
	(0.0002)	(0.0002)	(0.0002)
$NPV(p = 0.05)$	0.980	0.979	0.979
	(0.0000)	(0.0000)	(0.0000)
$TPR(p = 0.30)$	0.620	0.620	0.620
	(0.0008)	(0.0008)	(0.0008)
$FPR(p = 0.30)$	0.090	0.090	0.090
	(0.0001)	(0.0001)	(0.0001)
$NB(p = 0.30)$	0.071	0.071	0.071
	(0.0001)	(0.0001)	(0.0001)
$PPV(p = 0.30)$	0.578	0.579	0.579
	(0.0003)	(0.0004)	(0.0004)
$NPV(p = 0.30)$	0.923	0.923	0.923
	(0.0000)	(0.0000)	(0.0000)
$PCF(0.20)$	0.658	0.657	0.657
	(0.0004)	(0.0004)	(0.0004)
$PNF(0.85)$	0.623	0.623	0.623
	(0.0004)	(0.0004)	(0.0004)

TABLE 3

Efficiency comparisons of TIPW and AIPW estimators in a random sample design

(a) Random CCH designs with cohort sample size 1,000 and 1 control per case matching.

Method:		TIPW(δ, Y_1)		AIPW(δ, Y_1)		
	True	Mean	SE	Mean	SE	RE
β_1	1.099	1.116	(0.201)	1.123	(0.183)	(1.209)
β_2	0.693	0.702	(0.198)	0.700	(0.199)	(0.982)
AUC	0.875	0.877	(0.015)	0.878	(0.013)	(1.296)
DMR	0.357	0.362	(0.039)	0.361	(0.034)	(1.293)
$TPR(p = 0.05)$	0.948	0.948	(0.008)	0.948	(0.008)	(1.182)
$FPR(p = 0.05)$	0.516	0.508	(0.053)	0.507	(0.047)	(1.243)
$NB(p = 0.05)$	0.135	0.136	(0.013)	0.136	(0.013)	(1.008)
$PPV(p = 0.05)$	0.269	0.273	(0.020)	0.274	(0.019)	(1.167)
$NPV(p = 0.05)$	0.979	0.979	(0.002)	0.979	(0.002)	(1.299)
$TPR(p = 0.30)$	0.621	0.620	(0.042)	0.622	(0.039)	(1.137)
$FPR(p = 0.30)$	0.092	0.090	(0.015)	0.091	(0.014)	(1.162)
$NB(p = 0.30)$	0.071	0.072	(0.011)	0.072	(0.011)	(1.053)
$PPV(p = 0.30)$	0.574	0.582	(0.032)	0.581	(0.028)	(1.284)
$NPV(p = 0.30)$	0.923	0.923	(0.008)	0.923	(0.007)	(1.221)
$PCF(0.20)$	0.655	0.657	(0.031)	0.657	(0.027)	(1.311)
$PNF(0.85)$	0.621	0.623	(0.031)	0.623	(0.027)	(1.274)

(b) Random NCC designs with cohort sample size 2,000 and 1 control per case matching.

Method:		TIPW(δ, Y_1)		AIPW(δ, Y_1)		
	True	Mean	SE	Mean	SE	RE
β_1	1.099	1.188	(0.453)	1.236	(0.404)	(1.259)
β_2	0.693	0.757	(0.446)	0.743	(0.450)	(0.980)
AUC	0.883	0.889	(0.025)	0.892	(0.019)	(1.667)
DMR	0.167	0.193	(0.058)	0.197	(0.044)	(1.750)
$TPR(p = 0.05)$	0.933	0.926	(0.021)	0.929	(0.017)	(1.583)
$FPR(p = 0.05)$	0.420	0.396	(0.081)	0.394	(0.067)	(1.478)
$PPV(p = 0.05)$	0.057	0.063	(0.014)	0.063	(0.010)	(1.850)
$NPV(p = 0.05)$	0.997	0.997	(0.001)	0.997	(0.001)	(1.458)
$TPR(p = 0.30)$	0.785	0.783	(0.057)	0.791	(0.048)	(1.429)
$FPR(p = 0.30)$	0.190	0.183	(0.048)	0.186	(0.041)	(1.373)
$PPV(p = 0.30)$	0.101	0.111	(0.023)	0.110	(0.018)	(1.649)
$PCF(0.20)$	0.779	0.782	(0.057)	0.787	(0.044)	(1.634)
$PNF(0.85)$	0.720	0.727	(0.060)	0.733	(0.046)	(1.640)

TABLE 4
Efficiency comparisons under various two-phase designs.

(a) Efficiencies relative to TIPW estimators under CCH sampling with optimal sampling fraction

Setting	A: Random		B: Matched		C: Balanced	
Matching	none		Z_3		Z_3	
Method	TIPW	AIPW(δ, Y_{old})	TIPW	AIPW(δ, Y_{old})	TIPW	AIPW(δ, Y_{old})
β_1	0.73	0.97	1.04	1.08	0.83	0.87
β_2	0.90	0.88	1.10	1.13	0.87	0.88
AUC	0.61	1.17	1.01	1.25	0.66	0.88
DMR	0.64	1.22	1.02	1.34	0.60	0.89
$TPR(p = 0.05)$	1.01	1.10	0.54	0.73	0.76	0.85
$FPR(p = 0.05)$	0.72	1.13	0.79	1.09	0.83	1.04
$NB(p = 0.05)$	0.95	0.97	1.01	0.99	0.73	0.74
$PPV(p = 0.05)$	0.68	1.07	0.80	0.96	0.63	0.74
$NPV(p = 0.05)$	0.83	1.34	0.49	0.76	0.82	1.00
$TPR(p = 0.30)$	0.66	0.99	1.02	1.06	0.63	0.74
$FPR(p = 0.30)$	0.81	0.99	1.09	1.07	0.81	0.95
$NB(p = 0.30)$	0.82	1.04	1.00	1.16	0.62	0.74
$PPV(p = 0.30)$	0.69	1.18	1.05	1.55	0.57	0.93
$NPV(p = 0.30)$	0.67	1.06	1.04	1.13	0.79	0.94
$PCF(0.20)$	0.65	1.15	1.06	1.14	0.74	0.97
$PNF(0.85)$	0.62	1.12	0.98	1.11	0.76	0.92

(b) Efficiencies relative to TIPW estimators under NCC design with random sampling.

Setting	D: Random		E: Matched	
Matching	none		Z_3	
Method	TIPW	AIPW(δ, Y_{old})	TIPW	AIPW(X, δ, Y_{old})
β_1	1.00	1.30	1.53	1.63
β_2	1.00	0.98	1.58	1.61
AUC	1.00	1.30	0.99	1.23
DMR	1.00	1.45	1.45	1.65
$TPR(p = 0.01)$	1.00	1.24	0.61	1.05
$FPR(p = 0.01)$	1.00	1.24	0.55	1.10
$PPV(p = 0.01)$	1.00	1.31	1.18	1.43
$NPV(p = 0.01)$	1.00	1.36	0.44	0.54
$TPR(p = 0.03)$	1.00	1.24	1.16	1.34
$FPR(p = 0.03)$	1.00	1.27	0.68	1.03
$PPV(p = 0.03)$	1.00	1.31	1.52	1.70
$NPV(p = 0.03)$	1.00	1.24	0.62	1.00
$PCF(0.20)$	1.00	1.26	1.00	1.26
$PNF(0.85)$	1.00	1.28	0.90	1.23