Biostatistics (2012), **0**, 0, *pp.* 1–17
doi:10.1093/biostatistics/ROC˙under˙TDS˙2012-05-16˙supplementary

ROC Curve Estimation Under Test-Result-Dependent Sampling

XIAOFEI WANG, JUNLING MA, STEPHEN L. GEORGE

xiaofei.wang@duke.edu

1. Theoretical Results for Section 4

Lemma 1 Let $\mathcal{Y}, \mathcal{D}$ and $\mathcal{X}$ be the supports of Y, D and X. Assume that $n_0/n \to \rho_0 > 0$ and $n_k/n \to \rho_k \geqslant 0, k = 1, \cdots, K-1$ as $n \to \infty$. Under regularity conditions $n^{1/2}(\hat{\eta} - \eta)$ converges to a mean zero normal distribution in a neighborhood of η, and $n^{1/2}(\hat{G}(D, X) - G(D, X))$ converges to a mean zero Gaussian process on $\mathcal{D} \times \mathcal{X}$.

Proof: Similar to the proof of the theorem in Wang and Zhou (2006), we can prove $n^{1/2}(\hat{\eta} - \eta)$ is subject to a normal distribution as n tends to infinity. Since n_0/n and $n_k/n, k = 1, \cdots, K-1$ converge, n_K/n converges. Let $n_K/n \to \rho_K$ as $n \to \infty$. Under the regularity conditions for empirical likelihood (Qin and Lawless, 1994; Owen, 2001) and by use of the Taylor expansion we have

$$\begin{aligned} \hat{\eta} - \eta &= -\left[\frac{1}{n}\frac{\partial^2 l_p(\eta)}{\partial\eta\partial\eta^T}\right]^{-1} \frac{1}{n}\frac{\partial l_p(\eta)}{\partial\eta} + o_p(n^{-1/2}) \\ &\triangleq \sum_{s=0}^{K} \frac{1}{n_s} \sum_{i=1}^{n_s} \rho_s(-H_\eta)^{-1} grad_{si} + o_p(n^{-1/2}), \end{aligned} \tag{S1.1}$$

where $\frac{\partial l_p(\eta)}{\partial\eta}$, as the gradient vector of $l_p(\eta)$, has the form of $\sum_{s=0}^{K}\sum_{i=1}^{n_s} grad_{si}$, and $H_\eta = \lim_{n\to\infty} \frac{1}{n}\frac{\partial^2 l_p(\eta)}{\partial\eta\partial\eta^T}$. Rewriting the sum in (S1.1) as $1/n_0 \sum_{i=1}^{n_0} + 1/n_1 \sum_{i=1}^{n_1} + \cdots + 1/n_K \sum_{i=1}^{n_K}$, we apply the central limit theorem to each term to obtain $n^{1/2}(\hat{\eta} - \eta)$ subject to a normal distri-

bution as n tends to infinity. The asymptotic variance matrix of $n^{1/2}(\hat{\eta}-\eta)$ can be obtained as $H_\eta^{-1}\text{Cov}\left(\sum_{s=0}^K \rho_s grad_{si}\right) H_\eta^{-1}$.

Following the proof of Theorem 1 in Qin and Lawless (1994), we have

$$\begin{aligned}\hat{G}(d,x)-G(d,x) &= \frac{1}{n}\sum_{s=0}^{K}\sum_{i=1}^{n_s}\left[1-G(d,x)-(S^{-1}B)^T P_{si}\right] I(x_{si}\leqslant x, d_{si}=d)+o_p(n^{-1/2}) \\ &\triangleq \sum_{s=0}^{K}\frac{1}{n_s}\sum_{i=1}^{n_s}\rho_s\zeta_{si}^0+o_p(n^{-1/2}), \end{aligned} \tag{S1.2}$$

where $\zeta_{si}^0 = \left[1-G(d,x)-(S^{-1}B)^T P_{si}\right] I(x_{si}\leqslant x, d_{si}=d)$, $P_{si}=(\Pr(y_{si}\in C_1|d_{si},x_{si})-\theta_1,\cdots,\Pr(y_{si}\in C_{K-1}|d_{si},x_{si})-\theta_{K-1})^T$, $S=\lim_{n\to\infty} 1/n\sum_{s=0}^K\sum_{i=1}^{n_s} P_{si}P_{si}^T$, and $B=\lim_{n\to\infty} 1/n\sum_{s=0}^K\sum_{i=1}^{n_s} P_{si}$. Applying the central limit theorem, we have $n^{1/2}(\hat{G}(d,x)-G(d,x))$ a Gaussian process on $\mathcal{D}\times\mathcal{X}$. The asymptotic variance of $n^{1/2}(\hat{G}(d,x)-G(d,x))$ is $\sum_{s=0}^K\rho_s\text{var}\left(\zeta_{si}^0\right)$.

LEMMA 2 Under the conditions specified in Theorem 1, we have $n^{1/2}(\hat{S}_{1X}(c)-S_{1X}(c))$ and $n^{1/2}(\hat{S}_{0X}(c)-S_{0X}(c))$ each converges to a mean zero Gaussian process on $\mathcal{Y}$, where $\mathcal{Y}$ is the support of Y.

PROOF: Let n_D and $n_{\bar{D}}$ be the numbers of diseased and non-diseased subjects respectively, and $\gamma=\sum_{s=0}^K\rho_s\rho_s^D$. We have $n_D/n\to\gamma>0$ as $n\to\infty$.

Denote $\frac{\partial l_p(\eta)}{\partial\beta}=\sum_{s=0}^K\sum_{i=1}^{n_s} grad'_{si}$, and $H_\beta=\lim_{n\to\infty}\frac{1}{n}\frac{\partial^2 l_p(\eta)}{\partial\beta\partial\beta^T}$. By the functional delta method, we have

$$\begin{aligned} n^{1/2}\left[\hat{S}_{1X}(c)-S_{1X}(c)\right] &= \tfrac{\partial S_{1X}(c)}{\partial\beta}n^{1/2}(\hat{\beta}-\beta)+o_p(1) \\ &= \gamma^{-1/2}n_D^{-1/2}\tfrac{\partial S_{1X}(c)}{\partial\beta}\sum_{s=0}^K\sum_{i=1}^{n_s}\chi_{si}+o_p(1), \\ &= \gamma^{-1}\tfrac{\partial S_{1X}(c)}{\partial\beta}\sum_{s=0}^K(\rho_s\rho_s^D)^{1/2}(n_s^D)^{-1/2}\sum_{i=1}^{n_s}\chi_{si}+o_p(1), \end{aligned}$$

where $\chi_{si}=(-\gamma^{-1}H_\beta)^{-1}grad'_{si}I(d_{si}=1)$. Since $(n_s^D)^{-1/2}\sum_{i=1}^{n_s}\chi_{si}, s=0,\cdots,K$ are mutual independent, the conclusion follows by applying the central limit theorem to each part.

Let $\varpi_{si}^{(1)}=\frac{\partial S_{1X}(c)}{\partial\beta}\chi_{si}$, then the asymptotic variance of $n^{1/2}\left[\hat{S}_{1X}(c)-S_{1X}(c)\right]$ is

$$\sum_{s=0}^K\rho_s\rho_s^D\text{var}(\varpi_{si}^{(1)}).$$

For $\hat{S}_{0X}(c)$, let $\varpi_{si}^{(0)} = \frac{\partial S_{0X}(c)}{\partial \beta}(-\gamma^{-1}H_\beta)^{-1}grad'_{si}I(d_{si}=0)$. The conclusion for $\hat{S}_{0X}(c)$ can be similarly obtained from the equation that $n^{1/2}\left[\hat{S}_{0X}(c) - S_{0X}(c)\right] = (1-\gamma)^{-1}\sum_{s=0}^{K}(\rho_s\rho_s^D)^{1/2}(n_s^D)^{-1/2}\sum_{i=1}^{n_s}\varpi_{si}^{(0)} + o_p(1)$.

THEOREM 1 Assume $n_s^D/n_s \to \rho_s^D, s = 0, 1, \cdots, K$ with $0 < \rho_0^D < 1$ and $0 \leqslant \rho_1^D, \cdots, \rho_K^D < 1$ as $n \to \infty$. Under the conditions specified in Lemma 1, we have $n^{1/2}(\hat{ROC}_X(t) - ROC_X(t))$ converges to a mean zero Gaussian process on $(0,1)$ with variance $\Omega_1 + \Omega_0$ for any given t, where $\Omega_1 = \left[\frac{\partial S_{1X}(c)}{\partial \beta^T}\right] H_\beta^{-1}\sum_{s=0}^{K}\rho_s\rho_s^D\mathrm{var}(\varpi_{si}^{(1)})H_\beta^{-1}\left[\frac{\partial S_{1X}(c)}{\partial \beta}\right]$, and

$$\Omega_0 = \left[\left(\frac{\partial S_{0X}(v_t)}{\partial v_t}\right)^{-1}\frac{\partial S_{1X}(v_t)}{\partial v_t}\right]^2\left[\frac{\partial S_{0X}(c)}{\partial \beta^T}\right]H_\beta^{-1}\sum_{s=0}^{K}\rho_s(1-\rho_s^D)\mathrm{var}(\varpi_{si}^{(0)})H_\beta^{-1}\left[\frac{\partial S_{0X}(c)}{\partial \beta}\right].$$

PROOF: For any given $t \in (0,1)$, denote $v_t = S_{0X}^{-1}(t)$. By use of the functional delta method and its application to quantiles in van der Vaart (1998), we have

$$\hat{ROC}_X(t) - ROC_X(t) = \left[\hat{S}_{1X}(v_t) - S_{1X}(v_t)\right] - \frac{\partial S_{1X}(v_t)}{\partial v_t}\left(\frac{\partial S_{0X}(v_t)}{\partial v_t}\right)^{-1}\left[\hat{S}_{0X}(v_t) - S_{0X}(v_t)\right] + o_p(n^{-1/2}),$$

where $\frac{\partial S_{1X}(v_t)}{\partial v_t} = \frac{\partial \Phi(\sigma_1^{-1}(\beta_0+\beta_D+\beta_X^TX+\beta_{DX}^TX_D-c))}{\partial c}\mid_{c=v_t}$, $\frac{\partial S_{0X}(v_t)}{\partial v_t} = \frac{\partial \Phi(\sigma_0^{-1}(\beta_0+\beta_X^TX-c))}{\partial c}\mid_{c=v_t}$. It follows that $n^{1/2}\left[\hat{ROC}_X(t) - ROC(t)\right]$ converges to a mean zero Gaussian process. Since the diseased subjects and the non-diseased subjects are independent, the variance of $n^{1/2}\left[\hat{ROC}_X(t) - ROC_X(t)\right]$ at t is $\Omega_1 + \Omega_0$.

LEMMA 3 Under the condition of Theorem 1, we have $n^{1/2}(\hat{S}_1(c) - S_1(c))$ and $n^{1/2}(\hat{S}_0(c) - S_0(c))$ each converges to a mean zero Gaussian process on $\mathcal{Y}$.

PROOF: For simplicity of notation, denote $G(D=1,X)$ and $G(D=0,X)$ by G_1 and G_0 respectively, and their estimators by $\hat{G}_1$ and $\hat{G}_0$ respectively.

Since $S_1(c)$ is differentiable as a composite function of (η, G_1), applying the functional delta

method, we have

$$\begin{aligned} n^{1/2}\left[\hat{S}_1(c)-S_1(c)\right] &= n^{1/2}\left[\int \hat{S}_{1X}(c)\mathrm{d}\hat{G}_1/\int \mathrm{d}\hat{G}_1-\int S_{1X}(c)\mathrm{d}G_1/\int \mathrm{d}G_1\right] \\ &= (\int \mathrm{d}G_1)^{-1}\left[\int \frac{\partial S_{1X}(c)}{\partial \eta}\mathrm{d}G_1+\int S_{1X}(c)\frac{\partial \tilde{g}_1}{\partial \eta}\mathrm{d}x-S_1(c)\int \frac{\partial \tilde{g}_1}{\partial \eta}\mathrm{d}x\right] n^{1/2}(\hat{\eta}-\eta) \\ &\quad +n^{1/2}(\int \mathrm{d}G_1)^{-1}\left[\int (S_{1X}(c)-S_1(c))\mathrm{d}(\hat{G}_1-G_1)\right]+o_p(1), \end{aligned}$$

where $\tilde{g}_1=\frac{1}{n_0+\sum_{k=1}^K \frac{n_k}{\theta_k}\Pr(Y\in C_k|d,x)}I(d=1)$. It follows from Lemma 1 that $n^{1/2}\left[\hat{S}_1(c)-S_1(c)\right]$ converges to a Gaussian process with mean zero on $\mathcal{Y}$.

By use of (S1.1) and (S1.2), we can write the above equation as

$$\begin{aligned} n^{1/2}\left[\hat{S}_1(c)-S_1(c)\right] &= \gamma^{-1/2}(\int \mathrm{d}G_1)^{-1}\left[\int \frac{\partial S_{1X}(c)}{\partial \eta}\mathrm{d}G_1+\int S_{1X}(c)\frac{\partial \tilde{g}_1}{\partial \eta}\mathrm{d}x-S_1(c)\int \frac{\partial \tilde{g}_1}{\partial \eta}\mathrm{d}x\right] \\ &\quad n_D^{-1/2}\sum_{s=0}^K\sum_{i=1}^{n_s}(-\gamma^{-1}H_\eta)^{-1}grad_{si}I(d_{si=1}) \\ &\quad +\gamma^{-1/2}(\int \mathrm{d}G_1)^{-1}n_D^{-1/2}\sum_{s=0}^K\sum_{i=1}^{n_s}\{[S_{1X_{si}}(c)-S_1(c)]\,\zeta_{si}I(d_{si}=1)\}+o_p(1) \\ &\triangleq \gamma^{-1/2}n_D^{-1/2}\sum_{s=0}^K\sum_{i=1}^{n_s}\omega_{si}^{(1)}+o_p(1), \end{aligned}$$

where $grad_{si}$ is seen in the proof of Lemma 1, $\omega_{si}^{(1)}$ is seen in (4.10) and $\zeta_{si}=1-G(d_{si},x_{si})-(S^{-1}B)^TP_{si}$. Asymptotic variance of $n^{1/2}(\hat{S}_1(c)-S_1(c))$ at c is $\sum_{s=0}^K\rho_s\rho_s^D\mathrm{var}(\omega_{si}^{(1)})$.

THEOREM 2 Under the condition of Theorem 1, we have $n^{1/2}(\hat{ROC}(t)-ROC(t))$ converges to a mean zero Gaussian process on $(0,1)$ with variance $\Gamma_1+\Gamma_2$ for any given t, where $\Gamma_1=\sum_{s=0}^K\rho_s\rho_s^D\mathrm{var}(\omega_{si}^{(1)})$ and $\Gamma_2=\left[\left(\frac{\partial S_0(v_t)}{\partial v_t}\right)^{-1}\frac{\partial S_1(v_t)}{\partial v_t}\right]^2\sum_{s=0}^K\rho_s(1-\rho_s^D)\mathrm{var}(\omega_{si}^{(0)})$.

PROOF: For $\hat{S}_0(c)$, define $\omega_{si}^{(0)}$ by replacing $S_{1X}(c),S_1(c),G_1,n_D$ with $S_{0X}(c),S_0(c),G_0,n_{\bar{D}}$ respectively and setting $d_{si}=0$ in (4.10), then asymptotic normality of $n^{1/2}\left[\hat{S}_0(c)-S_0(c)\right]$ follows similarly from the equation that $n^{1/2}\left[\hat{S}_0(c)-S_0(c)\right]=(1-\gamma)^{-1/2}n_{\bar{D}}^{-1/2}\sum_{s=0}^K\sum_{i=1}^{n_s}\omega_{si}^{(0)}+o_p(1)$.

Similar to the proof of Theorem 1, $\hat{ROC}(t)$ can be shown to be a Gaussian process with mean 0 asymptotically. The variance of $n^{1/2}\left[\hat{ROC}(t)-ROC(t)\right]$ is

$$\sum_{s=0}^K\rho_s\rho_s^D\mathrm{var}(\omega_{si}^{(1)})+\left[\left(\frac{\partial S_0(v_t)}{\partial v_t}\right)^{-1}\frac{\partial S_1(v_t)}{\partial v_t}\right]^2\sum_{s=0}^K\rho_s(1-\rho_s^D)\mathrm{var}(\omega_{si}^{(0)}).$$

2. Supplementary Materials for Section 6

2.1 *Supplementary Figure and Table for Main Text*

[Insert Figure S1 around here]

[Insert Table S1 around here]

2.2 *Comparison of Three TDS Methods*

Under the TDS design, the empirical likelihood method SL_{TDS} and the weighted likelihood method WL_{TDS} are able to estimate covariate-specific ROC curve, while all three methods - SL_{TDS}, WL_{TDS} and the nonparametric method NP_{TDS} - are able to estimate covariate-independent ROC curve. Table S2 lists the ratio of mean square error (MSE) of SL_{TDS} (or NP_{TDS}) to that of WL_{TDS}. The WL_{TDS} is chosen to be the reference because it provides estimates for both covariate-specific $ROC_X(t)$ and covariate-independent $ROC(t)$. Under all scenarios, SL_{TDS} yields the smallest MSEs for both covariate-specific $ROC_X(t)$ and covariate-independent $ROC(t)$, while NP_{TDS} has the largest MSE in estimating covariate-independent $ROC(t)$. We conclude that SL_{TDS} is the most efficient among the three TDS methods, which is understandable as neither WL_{TDS} nor NP_{TDS} has used all available information in the data.

[Insert Table S2 around here]

2.3 *Choice of Number of Regions, Cutoff Points and Subject Allocation*

To maximize the potential efficiency gain of the TDS design in practice, one needs to make good choice for the number of regions, cutoff points and subject allocation. We conducted extensive simulations to evaluate the impact of number of regions, cutoff points and subject allocation on the precision of ROC estimation. We investigated the impact of the following factors: (i) the number of regions in which test result Y is divided, e.g. 2, 3 or 4 regions; (ii) the locations of cutoff points, which are defined by the distance from the center of test result, e.g. $(\mu_Y - a\sigma_Y, \mu_Y + a\sigma_Y)$

with $a = 0.5, 1.0, 1.5$ for 3-regions; (iii) the proportion of TDC subjects in the central region, e.g. n_2/n_{TDC} for 3-regions, and the proportion of TDC subjects in the right region, e.g. n_3/n_{TDC} for 3-regions; and (iv) the proportions of TDC subjects in the entire cohort, i.e. n_{TDC}/n. In all conditions, the same binormal ROC model has been used with the same number of subjects in the analysis cohort $n = 300$.

First, let's consider the cases that the test result is divided into 2, 3 and 4 regions according to the location of cutoff points. For 2-regions, three choices of the single cutoff point $\mu_Y + a\sigma_Y$ where $a = 0.5, 1.0, 1.5$ and several patterns of subject allocation were used, including oversampling or undersampling subjects in the right region and balanced allocation. Table S3 lists the simulated standard errors (SE) under these conditions.

[Insert Table S3 around here]

For 3-regions, three choices of the paired cutoff points $(\mu_Y - a\sigma_Y, \mu_Y + a\sigma_Y)$ where $a = 0.5, 1.0, 1.5$ and several patterns of subject allocation were used, including oversampling or undersampling subjects in the tailed regions and balanced allocation. Table S4 lists the simulated standard errors (SE) under these conditions.

[Insert Table S4 around here]

For 4-regions, three choices of the tripled cutoff points $(\mu_Y - a\sigma_Y, \mu_Y, \mu_Y + a\sigma_Y)$ where $a = 0.5, 1.0, 1.5$ and several patterns of subject allocation were used, including oversampling or undersampling subjects in the tailed regions and balanced allocation. Table S5 lists the simulated standard errors (SE) under these conditions.

[Insert Table S5 around here] [Insert Figures S2, S3, S4 around here]

In summary, there is no clear advantage in dividing test result into more than 3 regions. Holding other factors constant, 4-regions tends to yield worse performance than 3-regions, though

more regions give greater flexibility in allocating TDC subjects to different intervals of Y. In some combinations, 2-regions yields comparable performance to 3-regions, but 2-regions cannot selectively allocate subjects two tailed regions of test result, which limits its use in practice. More carefully, we studied the case of 3-regions. As seen in Table S4, setting the cutoff points far away from the center tends to improve the precision of ROC estimates, especially those corresponding to low false positive rate (e.g. t=0.1). Allocating more subjects to the tailed regions (e.g. balanced allocation or U-shaped allocation) increases the precision of ROC estimate across different values of $t = FPR$, especially those at low (e.g. t=0.1) and high (e.g. t=0.9) false positive rate (Figure S2). Allocating more subjects to the right region with high test results increases the precision of ROC estimation at low false positive rate (e.g. t=0.1), but it starts to lose efficiency when all TDC subjects are allocated to the right region (Figure S3). Allocating more subjects to the TDC tends to improve estimation efficiency, but the additional efficiency gain is small after the proportion of the TDC subjects in the entire study cohort is higher than 50% (Figure S4). These findings are consistent with our understanding on how the TDS design works to gain efficiency. When the test result has a bell-shaped distribution, oversampling subjects with very low and high test results will increase the representation of subjects who contain more information about ROC curve in the study cohort. In summary, if one has an interest in the performance of the biomarker over the entire ROC curve as well as the region that corresponds to low false positive rate, we recommend to have 50% TDS subjects in the overall study cohort, 3-regions of test result, cutoff points at $(\mu_Y - 1.5\sigma_Y, \mu_Y + 1.5\sigma_Y)$ and a balanced allocation of TDC subjects among the three regions. This combination may not be the one yielding the smallest standard errors, but it will achieve most of the potential efficiency gain and offer other advantages over some of the best combinations, such as a faster accrual of subjects to the two tailed regions since subjects with very low or high test results are less prevalent in the population. If one is only interested in the performance of the ROC curve corresponding to low false positive rate, a more extreme strategy

of subject allocation can be adopted, such as allocating the majority (e.g. 70%) TDC subjects to the right region. The recommendation is based on our simulations arising from the binormal ROC model. Since the ROC curve is invariant to strictly monotonic increasing transformation on Y, the recommended procedures should work as long as the test result can be transformed into a bell-shaped distribution.

2.4 *Robustness of the Binormal ROC Model*

To evaluate the performance of the proposed estimator SL_{TDS} under a non-normal ROC model, we generated the test result data using non-normal ROC models. These models are similar to (2.1), but they have error kernels other than $\epsilon \sim N(0,1)$. Two non-normal error kernels, skew normal and logistic distribution, were studied via simulation. The skew normal distribution generalizes the normal distribution to allow non-zero skewness (Azzalini, 1985). For skew normal kernel, we used location 0, scale 1 and shape γ where $\gamma = (0.1, 0.5, -0.5)$. Under the three shape values, the skew normal kernel ϵ has mean (0.0794, 0.3568, -0.3568), standard deviation (0.9968, 0.9342, 0.9342) and skewness (0.0002, 0.0239, -0.0239), respectively. Like normal distribution, the logistic distribution is symmetric around the mean, but it has thicker tails than a normal distribution. For logistic kernel, we used location 0 and scale 1, yielding mean 0 and standard deviation 1.8138. Other simulation setups are the same as Tables 1-3 in Section 6. Table S6 lists Bias, SE and 95% CP when the proposed method SL_{TDS} assuming binormal ROC model are used to estimate $ROC_X(t)$ with data generated from non-normal ROC models. When the true ROC model is not binormal and the propose method SL_{TDS} is still used to estimate ROC curve, the proposed TDS method performs reasonably well with small bias and reasonably calibrated variance estimate. The finding is consistent with the previous findings of several authors (Swets and Pckett, 1982; Hanley, 1996; Pepe, 2003) on the relative robustness of the binormal ROC model in related settings. Moreover, if the true distribution of test results conditional on D and X is non-normal,

standard diagnostic tools developed for linear regression can be used to identify the problem and find appropriate transformations to remedy the problem. For instance, a log transformation can be applied to test result with long tail skewed to the right. Since the ROC curve is invariant to strictly monotonic increasing transformation on test result, transformation itself will not mislead us about the true shape of ROC curve. Moreover, if the investigators have knowledge of the true distribution of test result, specific generalized linear model with non-identity link functions and non-normal response distribution can be developed using the same framework for normal linear model.

[Insert Table S6 around here]

References

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics* **12**, 171-178.

Hanley, J.A. (1996). The use of the 'binormal' model for parametric ROC analysis of quantitative diagnostic tests. *Statistics in Medicine*, **15**, 1575-1585.

Owen, A.B. (2001). *Empirical Likelihood*. Chapman and Hall, New York.

Pepe, M.S. (2003). *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford University Press, New York.

Qin, J. and Lawless, J.F. (1994). Empirical likelihood and general estimating equations. *Annals of Statistics* **22**, 300-325.

Swets, J.A. and Pickett, R.M. (1982). *Evaluation of diagnostic systems: Methods from signal detecting theory*. Academic Press, New York.

Van der Vaart, A.W. (1998). *Asymptotic Statistics*. Cambridge University Press, 1998.

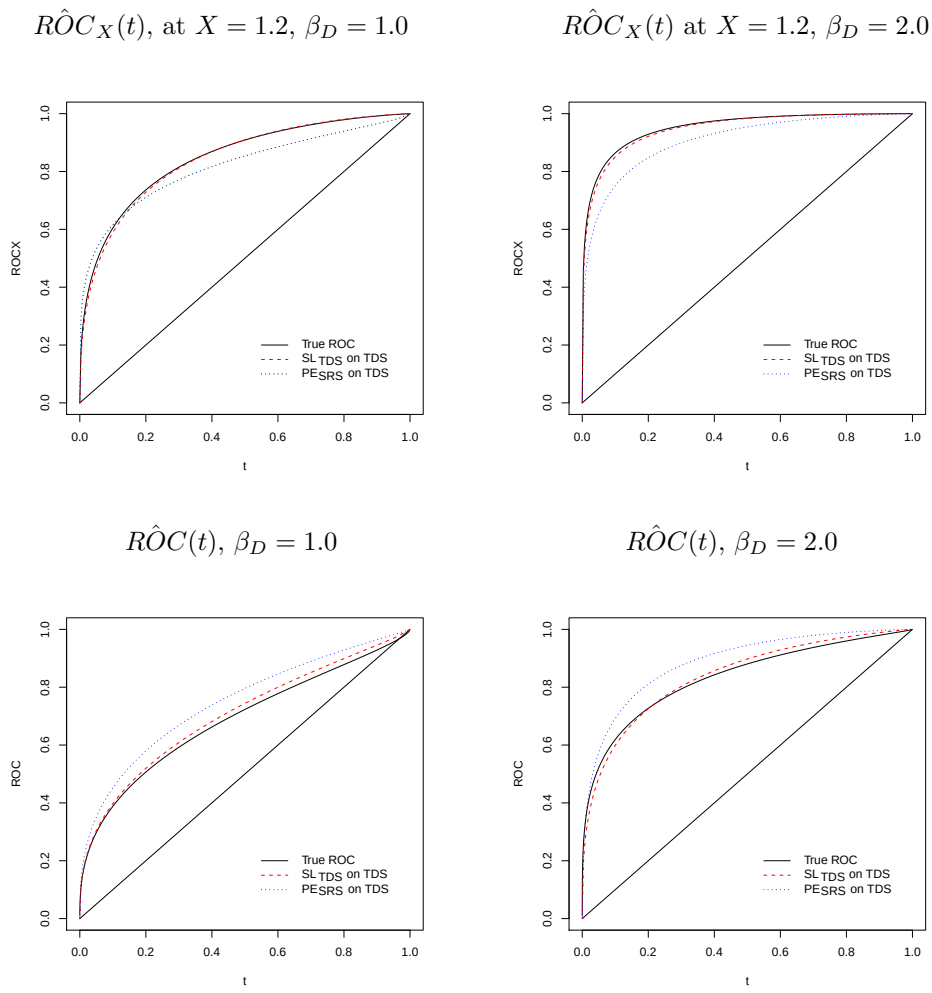

Fig. S1. ROC Curves Estimated by SL_{TDS} and PE_{SRS} on a TDS Dataset

True AUC_X=0.847 top left, 0.952 top right, 0.686 bottom left, 0.834 bottom right. SL_{TDS} $\hat{AUC}_X$=0.831 top left, 0.946 top right, 0.692 bottom left, 0.833 bottom right. PE_{SRS} $\hat{AUC}_X$=0.816 top left, 0.936 top right, 0.726 bottom left, 0.879 bottom right.

Table S1. Comparison of the TDS Methods (SL_{TDS}, WL_{TDS}) with IPW_{VB}

Method		True	Estimate	Bias%	SE	MSE	Ratio
			$\hat{ROC}_X(t)$ at $X = 0$				
SL_{TDS}	$t = 0.1$	0.407	0.409	0.49	0.0448	0.0020	0.89
	$t = 0.3$	0.654	0.657	0.46	0.0414	0.0017	0.91
	$t = 0.5$	0.798	0.800	0.25	0.0344	0.0012	0.92
	$t = 0.7$	0.898	0.899	0.11	0.0248	0.0006	0.94
WL_{TDS}	$t = 0.1$	0.407	0.408	0.25	0.0484	0.0023	1.04
	$t = 0.3$	0.654	0.657	0.46	0.0447	0.0020	1.06
	$t = 0.5$	0.798	0.800	0.25	0.0368	0.0014	1.06
	$t = 0.7$	0.898	0.899	0.11	0.0262	0.0007	1.05
IPW_{VB}	$t = 0.1$	0.407	0.409	0.49	0.0475	0.0023	1.00
	$t = 0.3$	0.654	0.657	0.44	0.0434	0.0019	1.00
	$t = 0.5$	0.798	0.801	0.46	0.0357	0.0013	1.00
	$t = 0.7$	0.898	0.900	0.38	0.0255	0.0007	1.00
			$\hat{ROC}(t)$				
SL_{TDS}	$t = 0.1$	0.385	0.386	0.52	0.0428	0.0018	1.00
	$t = 0.3$	0.588	0.592	0.68	0.0410	0.0017	1.06
	$t = 0.5$	0.724	0.723	-0.14	0.0369	0.0014	1.00
	$t = 0.7$	0.826	0.830	0.12	0.0300	0.0009	1.00
WL_{TDS}	$t = 0.1$	0.385	0.381	-0.78	0.0474	0.0023	1.28
	$t = 0.3$	0.588	0.595	1.19	0.0458	0.0021	1.31
	$t = 0.5$	0.724	0.731	0.97	0.0407	0.0017	1.21
	$t = 0.7$	0.829	0.839	1.21	0.0313	0.0011	1.22
IPW_{VB}	$t = 0.1$	0.385	0.382	-0.52	0.0426	0.0018	1.00
	$t = 0.3$	0.588	0.597	1.53	0.0395	0.0016	1.00
	$t = 0.5$	0.724	0.732	1.10	0.0366	0.0014	1.00
	$t = 0.7$	0.829	0.840	1.33	0.0280	0.0009	1.00

Table S2. Ratio of MSE(SL_{TDS}) (or MSE(NP_{TDS})) to MSE(WL_{TDS}) under TDS

	Method	$t=0.1$	$t=0.3$	$t=0.5$	$t=0.7$	$t=0.9$
$(a_1,a_2)=(\mu_Y-\sigma_Y,\mu_Y+\sigma_Y)$, $(n_0,n_1,n_2,n_3)=(150,50,50,50)$						
$\hat{ROC}_{X=0}(t)$	SL_{TDS}	0.88	0.86	0.80	0.67	0.33
$\hat{ROC}_{X=1.2}(t)$	SL_{TDS}	0.88	0.87	0.75	0.60	0.16
$\hat{ROC}(t)$	SL_{TDS}	0.83	0.82	0.82	0.83	1.00
$\hat{ROC}(t)$	NP_{TDS}	1.65	1.77	1.76	1.58	1.50
$(a_1,a_2)=(\mu_Y-\sigma_Y,\mu_Y+\sigma_Y)$, $(n_0,n_1,n_2,n_3)=(150,75,0,75)$						
$\hat{ROC}_{X=0}(t)$	SL_{TDS}	0.69	0.64	0.65	0.63	0.50
$\hat{ROC}_{X=1.2}(t)$	SL_{TDS}	0.68	0.67	0.67	0.75	0.28
$\hat{ROC}(t)$	SL_{TDS}	0.65	0.64	0.65	0.73	0.75
$\hat{ROC}(t)$	NP_{TDS}	1.87	2.27	2.18	1.91	1.25
$(a_1,a_2)=(\mu_Y-1.5\sigma_Y,\mu_Y+1.5\sigma_Y)$, $(n_0,n_1,n_2,n_3)=(150,75,0,75)$						
$\hat{ROC}_{X=0}(t)$	SL_{TDS}	0.49	0.45	0.48	0.50	0.50
$\hat{ROC}_{X=1.2}(t)$	SL_{TDS}	0.51	0.49	0.50	0.40	0.26
$\hat{ROC}(t)$	SL_{TDS}	0.42	0.38	0.38	0.40	0.50
$\hat{ROC}(t)$	NP_{TDS}	1.91	1.94	1.92	1.93	2.00

Table S3. Standard Errors of SL_{TDS} Estimates under 2-Regions

		$\hat{ROC}_X(t)$, $X=0$				$\hat{ROC}(t)$			
a	(n_0,n_1,n_2)	$t=0.1$	$t=0.3$	$t=0.5$	$t=0.7$	$t=0.1$	$t=0.3$	$t=0.5$	$t=0.7$
0.5	$(150,100,50)$	0.0512	0.0463	0.0381	0.0274	0.0488	0.0472	0.0427	0.0357
0.5	$(150,75,75)$	0.0493	0.0467	0.0392	0.0284	0.0468	0.0469	0.0433	0.0366
0.5	$(150,50,100)$	0.0476	0.0473	0.0407	0.0300	0.0448	0.0475	0.0451	0.0388
1.0	$(150,100,50)$	0.0479	0.0458	0.0385	0.0279	0.0453	0.0469	0.0435	0.0366
1.0	$(150,75,75)$	0.0462	0.0466	0.0402	0.0294	0.0439	0.0475	0.0450	0.0383
1.0	$(150,50,100)$	0.0468	0.0488	0.0422	0.0308	0.0436	0.0484	0.0464	0.0396
1.5	$(150,100,50)$	0.0469	0.0465	0.0394	0.0284	0.0441	0.0467	0.0437	0.0369
1.5	$(150,75,75)$	0.0458	0.0470	0.0403	0.0294	0.0433	0.0478	0.0454	0.0386
1.5	$(150,50,100)$	0.0457	0.0480	0.0416	0.0303	0.0425	0.0487	0.0470	0.0404

Table S4. Standard Errors of SL_{TDS} Estimates under 3-Regions

		$\hat{ROC}_X(t)$, $X=0$				$\hat{ROC}(t)$			
a	(n_0, n_1, n_2, n_3)	$t=0.1$	$t=0.3$	$t=0.5$	$t=0.7$	$t=0.1$	$t=0.3$	$t=0.5$	$t=0.7$
(-0.5, 0.5)	$(150, 100, 50, 0)$	0.0640	0.0525	0.0390	0.0260	0.0597	0.0530	0.0438	0.0339
(-0.5, 0.5)	$(150, 25, 100, 25)$	0.0563	0.0507	0.0418	0.0301	0.0559	0.0547	0.0491	0.0404
(-0.5, 0.5)	$(150, 0, 50, 100)$	0.0499	0.0510	0.0443	0.0328	0.0474	0.0526	0.0508	0.0439
(-0.5, 0.5)	$(150, 50, 50, 50)$	0.0503	0.0453	0.0371	0.0267	0.0479	0.0467	0.0423	0.0354
(-0.5, 0.5)	$(150, 75, 0, 75)$	0.0490	0.0438	0.0357	0.0254	0.0446	0.0426	0.0381	0.0317
(-1.0, 1.0)	$(150, 100, 50, 0)$	0.0626	0.0509	0.0380	0.0270	0.0592	0.0512	0.0412	0.0315
(-1.0, 1.0)	$(150, 25, 100, 25)$	0.0514	0.0473	0.0390	0.0279	0.0495	0.0481	0.0433	0.0358
(-1.0, 1.0)	$(150, 0, 50, 100)$	0.0468	0.0495	0.0431	0.0317	0.0449	0.0512	0.0497	0.0428
(-1.0, 1.0)	$(150, 50, 50, 50)$	0.0457	0.0417	0.0342	0.0244	0.0424	0.0407	0.0364	0.0302
(-1.0, 1.0)	$(150, 75, 0, 75)$	0.0438	0.0396	0.0323	0.0230	0.0386	0.0374	0.0335	0.0277
(-1.5, 1.5)	$(150, 100, 50, 0)$	0.0652	0.0528	0.0375	0.0233	0.0627	0.0527	0.0410	0.0307
(-1.5, 1.5)	$(150, 25, 100, 25)$	0.0488	0.0438	0.0352	0.0246	0.0465	0.0439	0.0381	0.0305
(-1.5, 1.5)	$(150, 50, 50, 50)$	0.0443	0.0399	0.0321	0.0224	0.0414	0.0392	0.0338	0.0266
(-1.5, 1.5)	$(150, 0, 50, 100)$	0.0478	0.0518	0.0452	0.0330	0.0433	0.0511	0.0499	0.0429
(-1.5, 1.5)	$(150, 75, 0, 75)$	0.0424	0.0383	0.0307	0.0214	0.0390	0.0363	0.0307	0.0238

Table S5. Standard Errors of SL_{TDS} Estimates under 4-Regions

		$\hat{ROC}_X(t)$, $X=0$				$\hat{ROC}(t)$			
a	$(n_0, n_1, n_2, n_3, n_4)$	$t=0.1$	$t=0.3$	$t=0.5$	$t=0.7$	$t=0.1$	$t=0.3$	$t=0.5$	$t=0.7$
(-0.5, 0, 0.5)	$(150, 10, 65, 65, 10)$	0.0600	0.0536	0.0456	0.0368	0.0594	0.0575	0.0527	0.0453
(-0.5, 0, 0.5)	$(150, 40, 35, 35, 40)$	0.0513	0.0467	0.0387	0.0279	0.0493	0.0477	0.0431	0.0359
(-0.5, 0, 0.5)	$(150, 70, 5, 5, 70)$	0.0473	0.0429	0.0350	0.0247	0.0439	0.0422	0.0382	0.0320
(-1.0, 0, 1.0)	$(150, 10, 65, 65, 10)$	0.0565	0.0499	0.0389	0.0263	0.0538	0.0509	0.0452	0.0372
(-1.0, 0, 1.0)	$(150, 40, 35, 35, 40)$	0.0546	0.0555	0.0482	0.0352	0.0506	0.0528	0.0522	0.0501
(-1.0, 0, 1.0)	$(150, 70, 5, 5, 70)$	0.0517	0.0465	0.0349	0.0228	0.0443	0.0388	0.0314	0.0250
(-1.5, 0, 1.5)	$(150, 10, 65, 65, 10)$	0.0532	0.0548	0.0469	0.0342	0.0523	0.0481	0.0419	0.0344
(-1.5, 0, 1.5)	$(150, 40, 35, 35, 40)$	0.0501	0.0424	0.0385	0.0323	0.0434	0.0412	0.0352	0.0273
(-1.5, 0, 1.5)	$(150, 70, 5, 5, 70)$	0.0431	0.0438	0.0440	0.0424	0.0345	0.0292	0.0306	0.0316

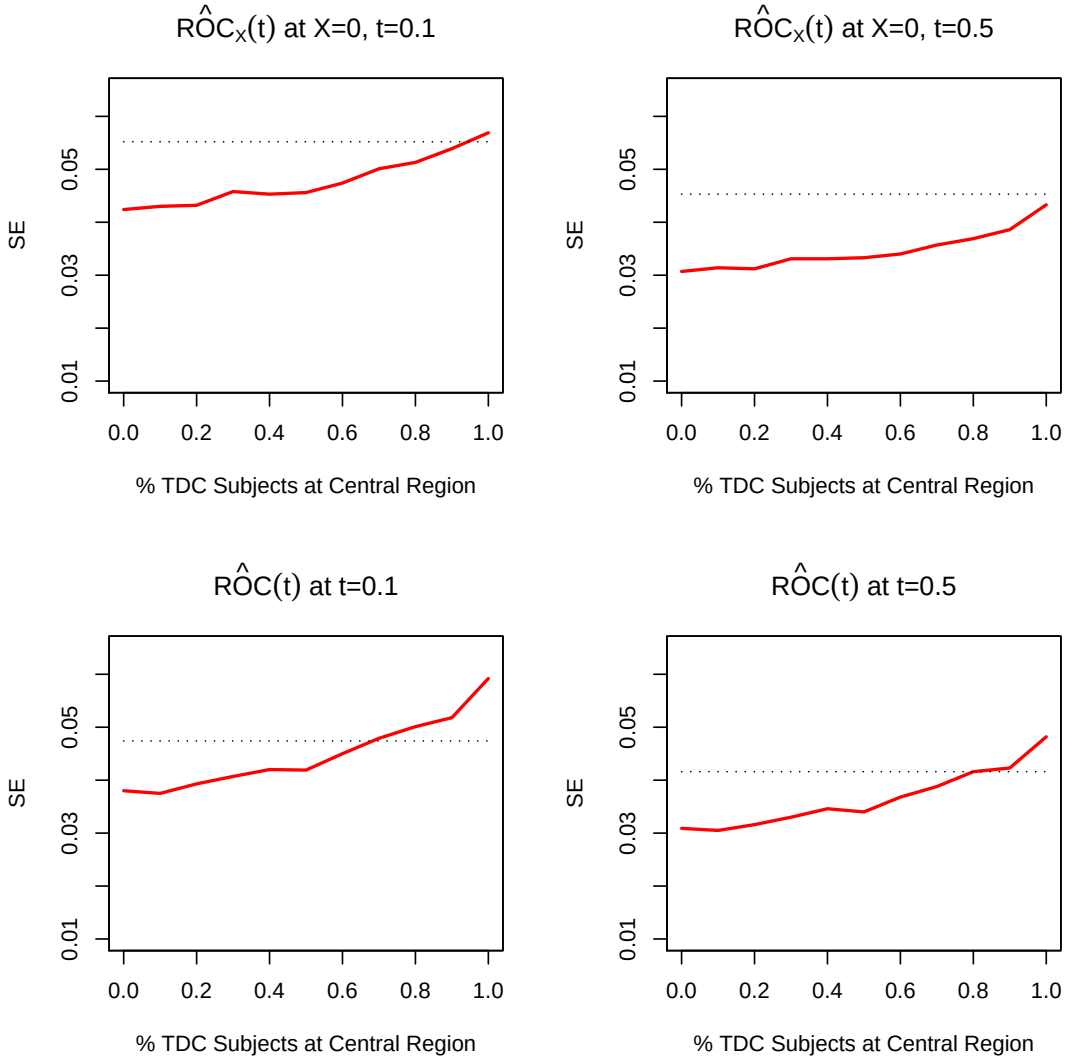

Fig. S2. Impact of % TDC Subjects in Central Region ($\frac{n_2}{n_{TDC}}$) on SE of ROC Estimates from SL_{TDS}

$n_{TDC} = 150, n_1 = n_3 = (n_{TDC} - n_2)/2, n_0 = 150$. Dotted line is SE from PE^*_{SRS} in Section 6.4.

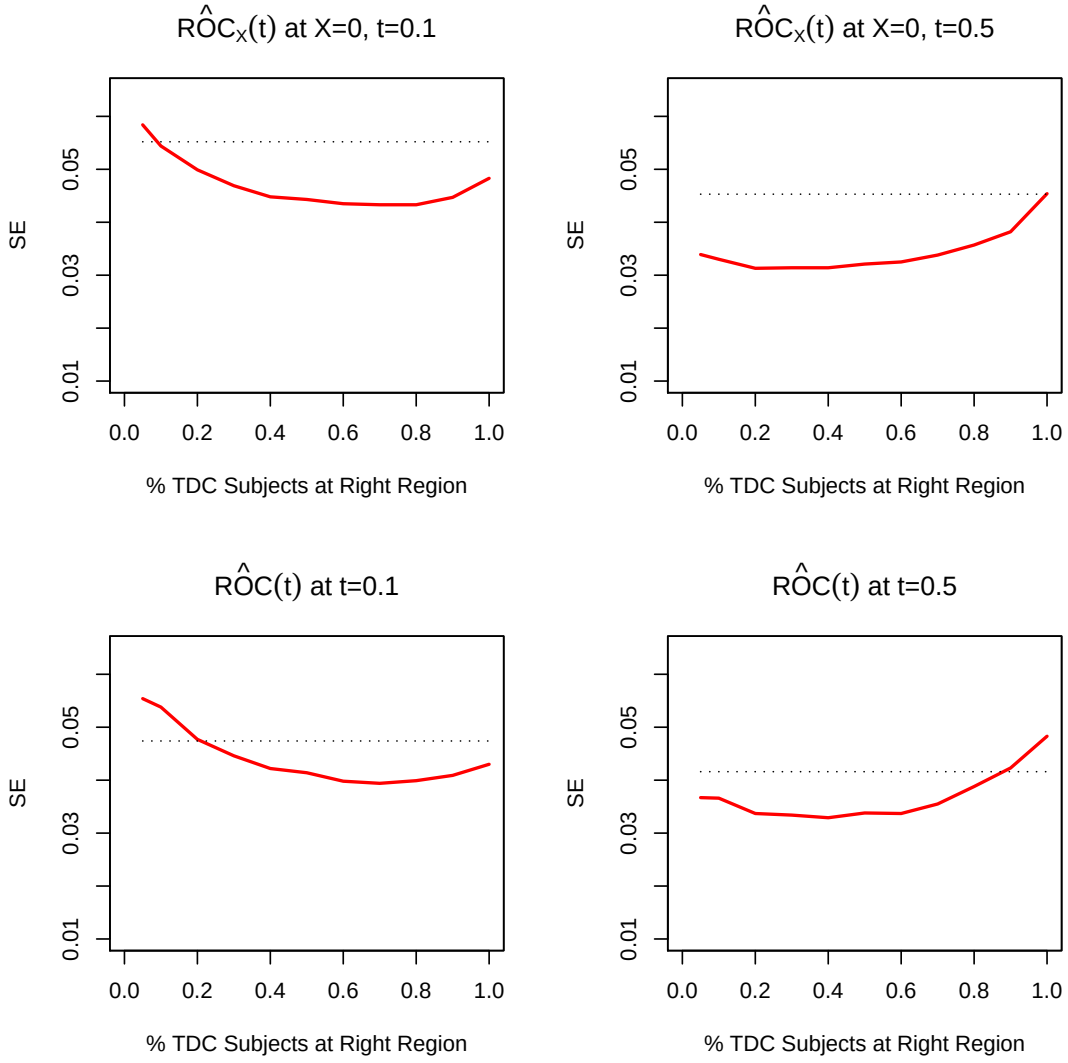

Fig. S3. Impact of % TDC Subjects in Right Region ($\frac{n_3}{n_{TDC}}$) on SE of ROC Estimates from SL_{TDS} $n_{TDC} = 150, n_2 = 50, n_1 = n_{TDC} - n_2 - n_3, n_0 = 150$. Dotted line is SE from PE^*_{SRS} in Section 6.4.

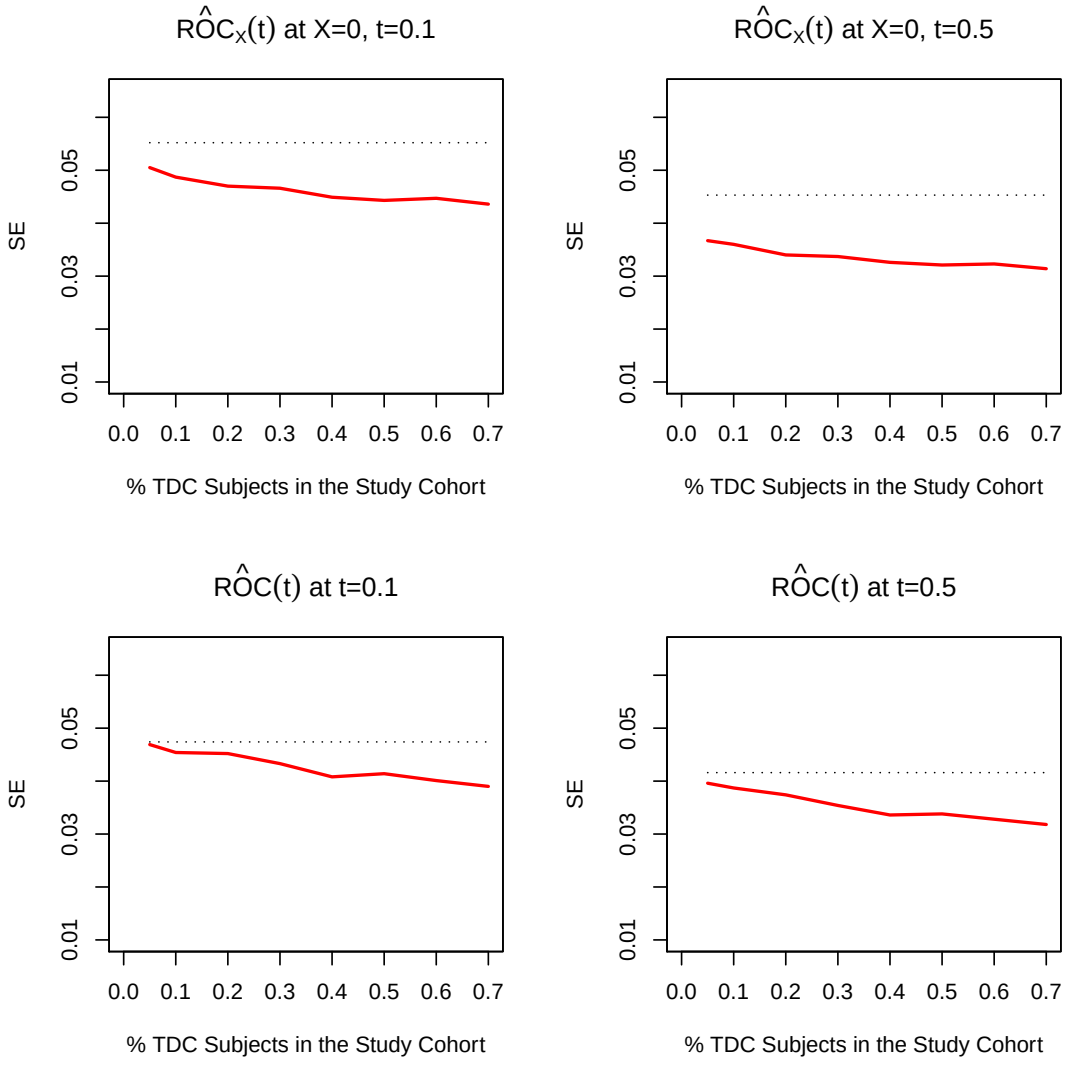

Fig. S4. Impact of % TDC Subjects in the Study Cohort ($\frac{n_{TDC}}{n}$) on SE of ROC Estimates from SL_{TDS}

$n = 300, n_1 = n_2 = n_3 = n_{TDC}/3, n_0 = n - n_{TDC}$. Dotted line is SE from PE^*_{SRS} in Section 6.4.

Table S6. $ROC_X(t)$ at $X = 0$ Estimated by SL_{TDS} under Non-Normal Error Kernels

		True	Estimate	Bias%	SE	$\hat{\text{SE}}$	95% CP
skew normal	$t = 0.1$	0.413	0.417	0.87	0.0466	0.0531	0.956
$\gamma = 0.1$	$t = 0.3$	0.660	0.663	0.47	0.0428	0.0519	0.972
	$t = 0.5$	0.802	0.804	0.24	0.0351	0.0420	0.964
	$t = 0.7$	0.901	0.902	0.13	0.0250	0.0287	0.950
skew normal	$t = 0.1$	0.453	0.462	1.99	0.0686	0.0825	0.942
$\gamma = 0.5$	$t = 0.3$	0.698	0.703	0.70	0.0505	0.0825	0.962
	$t = 0.5$	0.831	0.833	0.20	0.0374	0.0650	0.962
	$t = 0.7$	0.919	0.919	-0.01	0.0247	0.0422	0.962
skew normal	$t = 0.1$	0.408	0.406	-0.42	0.0477	0.1274	0.979
$\gamma = -0.5$	$t = 0.3$	0.653	0.652	-0.12	0.0441	0.1164	0.965
	$t = 0.5$	0.796	0.796	0.03	0.0358	0.0877	0.963
	$t = 0.7$	0.896	0.896	-0.03	0.0253	0.0555	0.941
logistic	$t = 0.1$	0.269	0.263	-2.38	0.0450	0.0386	0.908
	$t = 0.3$	0.532	0.503	-5.42	0.0503	0.0434	0.909
	$t = 0.5$	0.697	0.674	-3.31	0.0519	0.0403	0.914
	$t = 0.7$	0.823	0.816	-0.90	0.0384	0.0327	0.921