1-21-2011

Semiparametric Estimation of the Covariate-Specific ROC Curve in Presence of Ignorable Verification Bias

Danping Liu
University of Washington - Seattle Campus, dpliu@u.washington.edu

Xiao-Hua Zhou
University of Washington, azhou@u.washington.edu

Semiparametric estimation of the covariate-specific ROC curve in presence of ignorable verification bias

Danping Liu*

Department of Biostatistics, University of Washington, Seattle, WA 98195

**email:* dpliu@u.washington.edu

and

Xiao-Hua Zhou*

Northwest HSR&D Center of Excellence, VA Puget Sound Health Care System, Seattle, WA 98108

Department of Biostatistics, University of Washington, Seattle, WA 98195

**email:* azhou@u.washington.edu

SUMMARY: Covariate-specific ROC curves are often used to evaluate the classification accuracy of a medical diagnostic test or a biomarker, when the accuracy of the test is associated with certain covariates. In many large-scale screening tests, the gold standard is subject to missingness due to high cost or harmfulness to the patient. In this paper, we propose a semiparametric estimation of the covariate-specific ROC curves with a partial missing gold standard. A location-scale model is constructed for the test result to model the covariates' effect, but the residual distributions are left unspecified. Thus the baseline and link functions of the ROC curve both have flexible shapes. With the gold standard missing at random (MAR) assumption, we consider weighted estimating equations for the location-scale parameters, and weighted kernel estimating equations for the residual distributions. Three ROC curve estimators are proposed and compared, namely, imputation-based, inverse probability weighted and doubly robust estimators. We derive the asymptotic normality of the estimated ROC curve, as well as the analytical form the standard error estimator. The proposed method is motivated and applied to the data in an Alzheimer's disease research.

KEY WORDS: Alzheimer's disease; covariate-specific ROC curve; ignorable missingness; verification bias; weighted estimating equations.

1

1. Introduction

The receiver operating characteristic (ROC) curve is a useful tool to evaluate the classification ability of a medical diagnostic test or biomarker. The ROC curve is a plot of test's sensitivity versus its 1-specificity as one varies the decision threshold for test positivity. In the ROC analysis, covariates may impact the magnitude of the diagnostic test and/or the diagnostic accuracy. Lack of covariate adjustment may not only bias the result, but also impair the generalizability of the study results to other different populations. Thus the covariate-specific ROC curve is widely used to evaluate the classification accuracy within some particular sub-population. One may consider a stratified analysis and estimate the ROC curve for each sub-population specified by the covariates. However, regression type analysis is often preferred so that the covariates' effect is estimated in a parsimonious fashion. Zhou *et al.* (2002) and Pepe (2003) both give a detailed review of the existing methods in estimating a covariate-specific ROC curve. Many methods require the true condition status of each patient to be determined by the "gold standard". In many large cohort studies, however, the gold standard may not be available to everybody because it is expensive and/or invasive. Deleting the subjects with missing gold results in biased estimators, which is referred to as the "verification bias" (Begg and Greenes, 1983).

Using the missing data framework, we call the verification process to be missing at random (MAR) if the probability of disease verification is only affected by the observed variables. Under the MAR assumption, many existing methods are available for the verification bias problem for binary tests (Begg and Greenes, 1983) and ordinal tests (Gary *et al.*, 1984; Zhou, 1996; Zhou, 1998; Rodenberg and Zhou, 2000). Recently, Zheng *el al.* (2005) proposed to use a weighted estimating equation (WEE) approach to estimate the covariate specific ROC curve for ordinal tests. They considered a parametric binormal form of the ROC curve. The theory of their WEE approach originated from Lipsitz *et al.* (1999). The weighted estimating

equations utilize a model of missing mechanism as well as a model of disease probability in order to estimate the parameters in the binormal ROC curve. The advantage of their method is the robustness to some model mis-specification: either the disease model or the verification model being correct guarantees the consistency of the ROC curve estimator, which is called "doubly robust" property. Alonzo and Pepe (2005) considered the continuous test result and proposed several empirical ROC curve estimators. Their ROC curve estimators are empirical step functions, and could not incorporate the covariates effect on ROC curves.

Although covariate-specific ROC curve estimators for continuous test have been extensively discussed in the literature, not much work has been done on the verification bias correction for covariate-specific ROC curves. Page and Rotnitzky (2009) is the only published paper so far, who proposed a fully parametric model for estimating the covariate-specific ROC curve under verification bias. However, their binormal ROC curve assumption is often too restrictive in practice, so we wish to estimate the "baseline shape" of the ROC curve instead. A subgroup analysis is an option when the covariates are categorized. For example, Punglia *et al.* (2003) studied the ROC curve for prostate-specific antegen (PSA) measurement in detecting prostate cancer. They reported the bias corrected ROC curve stratified by age group and digital rectal examination results. When some of the covariates are continuous, however, subgroup analysis may not be feasible. In this paper, we propose a new semiparametric regression model for the covariate-specific ROC curve and the weighted estimating equations to adjust for the verification bias, which extends the results in Alonzo and Pepe (2005) and Page and Rotnitzky (2009).

We consider a continuous-scale diagnostic test and propose several semiparametric covariate-specific ROC curve estimators. A location-scale model is constructed on the diagnostic test to model the covariates effect, but the residual distributions are left unspecified. This location-scale framework is commonly used in regression settings as well. The baseline and

link function of the ROC curve both have flexible shapes. Pepe (1998) first proposed and compared several regression methods to estimate the ROC curve without missing data, and showed that the location-scale model is the most efficient. With missing gold standard, we employ the weighted estimating equations for the location-scale parameters, similarly as in Zheng *el al.* (2005) and Page and Rotnitzky (2009). The unspecified residual distributions are estimated by the weighted kernel estimating equations, which yields the smooth ROC curve estimators. We discuss three forms of weighting techniques based on imputation and inverse probability weighting. The covariate-specific ROC curve is then estimated as a function of the location-scale parameters and the residual distribution/quantile functions. We also show the central limit theorem for the estimated ROC curve and derive the asymptotic variance formula. Compared to Alonzo and Pepe (2005), our approach can incorporate covariates, and we derive analytical variance formula; compared to Zheng *el al.* (2005) and Page and Rotnitzky (2009), the form of our ROC curve is more flexible, as we do not specify the baseline shape of the curve.

The paper is organized as follows. Section 2 outlines the location-scale model framework and other basic model assumptions. Section 3 presents the weighted estimating equations for the finite and infinite dimensional parameters. The limit theorems are also presented. Section 4 reports some simulation results to examine the finite sample performance of our proposed method. Section 5 applies our method to a real data example in an Alzheimer's disease research, followed by discussions in Section 6.

2. Location-scale model

Let T_i, D_i, V_i and X_i be the continuous test result, the gold standard, the verification indicator, and the covariates for the ith subject, respectively, where $i = 1, 2, \cdots, n$. Let a larger value of T indicate more likely to be diseased; let $D = 1$ denote a diseased subject (case) and 0 denote a healthy subject (control); let $V = 1$ denote observed gold standard

and 0 denote missing gold standard. We sometimes suppress the subscript i when there is no confusion.

We assume a location-scale model for T_i:

$$T_i = \mu(X_i, D_i; \beta) + \sigma(X_i, D_i; \gamma) \times \epsilon_i(D_i),$$

where $\mu(X_i, D_i; \beta)$ and $\sigma(X_i, D_i; \gamma)$ are the mean and standard deviation for T_i given the values of the covariates and disease status, respectively, and $\epsilon_i(D_i)$ is the residual. We may also use μ_i and σ_i as an abbreviation. Let G_0 and G_1 be the unknown distribution functions for $\epsilon_i(0)$ and $\epsilon_i(1)$, respectively, with mean 0 and variance 1. Our setting of the location-scale model slightly extended the model in Section 2 of Pepe (1998), in that we allow the two distribution functions to be different. Page and Rotnitzky (2009) also used this location-scale model, but assumed the residual $\epsilon_i(D_i)$ to follow standard normal distribution. As we will see in the example in Section 5, the cases and controls can have quite different test distributions, both of which are far from normal.

We could then write out the covariate-specific sensitivity and specificity at some cutoff point c:

$$
\begin{aligned}
Sens_x(c) &= \Pr(T \geqslant c \mid D = 1, X = x) \\
&= 1 - G_1\left(\frac{c - \mu(x, 1; \beta)}{\sigma(x, 1; \gamma)}\right), \\
1 - Spec_x(c) &= \Pr(T \geqslant c \mid D = 0, X = x) \\
&= 1 - G_0\left(\frac{c - \mu(x, 0; \beta)}{\sigma(x, 0; \gamma)}\right).
\end{aligned}
$$

The ROC curve is expressed as follows:

$$ROC_x(t) = 1 - G_1\left[\frac{\sigma(x, 0; \gamma)}{\sigma(x, 1; \gamma)} G_0^{-1}(1 - t) + \frac{\mu(x, 0; \beta) - \mu(x, 1; \beta)}{\sigma(x, 1; \gamma)}\right].$$

Sometimes we call G_1 and G_0^{-1} the link function and the baseline of the ROC curve, respectively. Allowing the link and baseline to be unspecified grants the ROC curve more flexibility. As a comparison, Pepe (1998), and Cai and Pepe (2002) both proposed a direct

regression method for estimating the covariate-specific ROC curve; the former assumes both G_1 and G_0^{-1} have known parametric forms, while the latter assumes G_1 to be known and leaves G_0^{-1} to be unspecified. Pepe (1998) also mentioned that the location-scale model tends to be more efficient than the direct regression method. Another advantage of the location-scale model is that its extension to missing gold standard situation is more straightforward.

3. Estimation procedures

3.1 *Complete data estimating equations*

In the ROC curve expression, the unknown quantities to be estimated are finite dimensional parameters β and γ, and infinite dimensional curves G_0^{-1} and G_1. When the gold standard is observed for each subject, the estimation of β and γ is easily obtained via the following estimating functions:

$$U_1 \equiv \sum_i U_{1i} \equiv \sum_i \left(\frac{\partial \mu_i}{\partial \beta} \right)^T \frac{(T_i - \mu_i)}{\sigma_i^2},$$

$$U_2 \equiv \sum_i U_{2i} \equiv \sum_i \left(\frac{\partial}{\partial \gamma} \sigma_i^2 \right)^T \frac{(T_i - \mu_i)^2 - \sigma_i^2}{\operatorname{var}\left[(T_i - \mu_i)^2 \right]},$$

where μ_i and σ_i are short for $\mu(X_i, D_i; \beta)$ and $\sigma(X_i, D_i; \gamma)$ defined in the location-scale model. Substituting the estimated $\hat{\beta}$ and $\hat{\gamma}$ into μ_i and σ_i, we denote $\hat{\epsilon}_i(D_i) = \frac{(T_i - \hat{\mu}_i)}{\hat{\sigma}_i}$ to be the fitted residual. For each fixed s, the two distribution functions, $G_1(s)$ and $G_0(s)$, can then be estimated with the following kernel smoothing estimators:

$$U_3(s) \equiv \sum_i U_{3i}(s) \equiv \sum_i D_i \left[K \left(\frac{s - \hat{\epsilon}_i(D_i)}{h} \right) - G_1(s) \right],$$

$$U_4(s) \equiv \sum_i U_{4i}(s) \equiv \sum_i (1 - D_i) \left[K \left(\frac{s - \hat{\epsilon}_i(D_i)}{h} \right) - G_0(s) \right],$$

where $K(\cdot)$ is some distribution function and h is the bandwidth. In order to obtain the $1 - t$ quantile of G_0, $G_0^{-1}(1 - t)$, we solve the following estimating function for every fixed t:

$$U_5(t) \equiv \sum_i U_{5i}(t) \equiv \sum_i (1 - D_i) \left[K \left(\frac{G_0^{-1}(1 - t) - \hat{\epsilon}_i(0)}{h} \right) + t - 1 \right].$$

Although t may take infinitely many values between 0 and 1, we set finite grid points to get good approximation for the smooth ROC curve. In the example, we choose $t = 0.01, 0.02, \cdots, 0.99$ with linear extrapolation between the adjacent grid points. When the kernel $K(x) = I(x \leqslant 0)$, we obtain the empirical estimators of G_0 and G_1. However, it is usually desired that the ROC curve is a smooth curve rather than a step function. Therefore, we use some continuous distribution function $K(\cdot)$, such as standard normal distribution Φ. The estimated $\hat{G}_0$ and $\hat{G}_1$ have the same $\sqrt{n}$-consistency as the empirical distribution function, as long as the bandwidth is sufficiently small (Nadaraya, 1964).

3.2 *Weighted estimating equations under verification bias*

As the gold standard is only available for a portion of the subjects, we reweight the estimating equations. Let $\rho_i = \Pr\left(D_i = 1 | T_i, X_i\right) = \Pr\left(D_i = 1 | T_i, X_i, V_i = 1\right)$ be the disease probability, and $\pi_i = \Pr\left(V_i = 1 | T_i, X_i\right)$ be the verification probability. With the MAR assumption, the two probabilities can be estimated separately, where logistic regressions would be a convenient approach. We then construct the weighted estimating equations with the estimated $\hat{\rho}_i$ and $\hat{\pi}_i$. Three types of estimating methods are considered, namely, doubly robust (DR), inverse probability weighting (IPW) and imputation based (IB) approaches.

Let

$$S_k^{DR} \equiv \sum_i S_{ki}^{DR} \equiv \sum_i \left\{ \frac{V_i}{\hat{\pi}_i} U_{ki} + \left(1 - \frac{V_i}{\hat{\pi}_i}\right) E_{D_i|T_i,X_i} U_{ki} \right\}, \tag{1}$$

$$S_k^{IPW} \equiv \sum_i S_{ki}^{IPW} \equiv \sum_i \frac{V_i}{\hat{\pi}_i} U_{ki}, \tag{2}$$

$$S_k^{IB} \equiv \sum_i S_{ki}^{IB} \equiv \sum_i \left\{ V_i U_{ki} + (1 - V_i) E_{D_i|T_i,X_i} U_{ki} \right\}, \tag{3}$$

for $k = 1, 2, 3, 4, 5$. We abbreviate the superscript and use S_k to denote the general weighted estimating functions in the text below. The conditional expectation $E_{D_i|T_i,X_i}$ can be written

as weighted summations, as D_i takes the value of 0 or 1. For example,

$$
\begin{aligned}
E_{D_i|T_i,X_i} U_{1i} &= \sum_{d=0}^{1} \Pr(D_i = d | T_i, X_i) \left(\frac{\partial \mu(X_i, d; \beta)}{\partial \beta} \right)^T \frac{(T_i - \mu(X_i, d; \beta))}{\sigma^2(X_i, d; \gamma)} \\
&= (1 - \rho_i) \left(\frac{\partial \mu(X_i, 0; \beta)}{\partial \beta} \right)^T \frac{(T_i - \mu(X_i, 0; \beta))}{\sigma^2(X_i, 0; \gamma)} \\
&\quad + \rho_i \left(\frac{\partial \mu(X_i, 1; \beta)}{\partial \beta} \right)^T \frac{(T_i - \mu(X_i, 1; \beta))}{\sigma^2(X_i, 1; \gamma)}.
\end{aligned}
$$

The DR estimating functions (1) enjoy the "doubly robust" property: as long as either $\hat{\rho}_i$ or $\hat{\pi}_i$ is consistently estimated, the DR estimator is consistent; the IPW estimating equations (2) require that the verification probability, π_i, is consistently estimated; the IB estimating equations (3) require that the disease probability, ρ_i, is consistently estimated. In practice, our understanding of the missing mechanism or the disease risk may not be accurate enough, so the DR estimator allows two shots for the model specification, while IPW and IB estimators rely on a single model assumption.

All these weighted estimating equations can be solved by Newton-Raphson method. With the estimated $\hat{\beta}$, $\hat{\gamma}$, $\hat{G}_0^{-1}$ and $\hat{G}_1$, we could estimate the covariate-specific ROC curve as follows:

$$
\widehat{ROC}_x(t) = 1 - \hat{G}_1 \left[\frac{\sigma(x, 0; \hat{\gamma})}{\sigma(x, 1; \hat{\gamma})} \hat{G}_0^{-1}(1 - t) + \frac{\mu(x, 0; \hat{\beta}) - \mu(x, 1; \hat{\beta})}{\sigma(x, 1; \hat{\gamma})} \right].
$$

3.3 *Asymptotic normality*

First we examine the asymptotic behavior of the location-scale parameters. Let $B_1(\alpha_1) = \sum_i B_{1i}(\alpha_1)$ and $B_2(\alpha_2) = \sum_i B_{2i}(\alpha_2)$ be the estimating functions for modelling ρ and π, respectively. Let $\theta_1 \equiv \begin{pmatrix} \beta \\ \gamma \end{pmatrix}$ be the location-scale parameters, and $\hat{\theta}_1$ be its estimated version. Let $U_{12,i} \equiv \begin{pmatrix} U_{1i} \\ U_{2i} \end{pmatrix}$ and $S_{12,i} \equiv \begin{pmatrix} S_{1i} \\ S_{2i} \end{pmatrix}$. Define $I \equiv -\frac{\partial}{\partial \theta_1^T} ES_{12,i}$, $J_1 = -\frac{\partial}{\partial \alpha_1^T} ES_{12,i}$, $J_2 = -\frac{\partial}{\partial \alpha_2^T} ES_{12,i}$, $K_1 = -\frac{\partial}{\partial \alpha_1^T} EB_{1i}$, $K_2 = -\frac{\partial}{\partial \alpha_2^T} EB_{2i}$. Let $Q_i = U_{12,i} - J_1 K_1^{-1} B_{1i} - J_2 K_2^{-1} B_{2i}$. The asymptotic distribution of $\hat{\theta}_1$ is stated in the following theorem:

THEOREM 1: *Under the standard regularity conditions stated in the Web Appendix,*

$$\sqrt{n}\left(\hat{\theta}_1 - \theta_1\right) \xrightarrow{d} N(0, \Omega_1),$$

where

$$\Omega_1 = I^{-1}\text{var}(Q_i)I^{-1}.$$

The proof of this theorem is sketched in the Web Appendix. Note that Ω_1 can be estimated by replacing all the parameters θ, α_1 and α_2 with their estimates, and replacing all the expectations in I, J_1, J_2, K_1 and K_2 with the sample mean.

Our primary interest is not to estimate the location and scale model, but to construct the ROC curve. Before studying the asymptotic property of the estimated ROC curve, we first take a look at the variances for $\hat{G}_1$ and $\hat{G}_0^{-1}$. The influence functions of $\hat{G}_0^{-1}(1-t)$ and $\hat{G}_1(s)$ are stated in the following lemma.

LEMMA 1: *When $n \to +\infty$, $h \to 0$ and $nh^4 \to 0$, both $\hat{G}_0^{-1}(1-t)$ and $\hat{G}_1(s)$ are asymptotic linear for any $t \in (0,1)$ and $s \in (-\infty, +\infty)$, :*

$$\sqrt{n}\left(\hat{G}_0^{-1}(1-t) - G_0^{-1}(1-t)\right) = \frac{1}{\sqrt{n}}\sum_i A_{1i}(t) + o_p(1),$$

$$\sqrt{n}\left(\hat{G}_1(s) - G_1(s)\right) = \frac{1}{\sqrt{n}}\sum_i A_{2i}(s) + o_p(1)$$

where

$$A_{1i}(t) = \left(-\frac{\partial}{\partial s}ES_{5i}\right)^{-1}\left[S_{5i} + \left(\frac{\partial}{\partial\theta_1^T}ES_{5i}\right)I^{-1}Q_i + \left(\frac{\partial}{\partial\alpha_1^T}ES_{5i}\right)K_1^{-1}B_{1i}\right.$$
$$\left. + \left(\frac{\partial}{\partial\alpha_2^T}ES_{5i}\right)K_2^{-1}B_{2i}\right]$$

, and

$$A_{2i}(s) = \left(-\frac{\partial}{\partial G_1(s)}ES_{3i}\right)^{-1}\left[S_{3i} + \left(\frac{\partial}{\partial\theta_1^T}ES_{3i}\right)I^{-1}Q_i + \left(\frac{\partial}{\partial\alpha_1^T}ES_{3i}\right)K_1^{-1}B_{1i}\right.$$
$$\left. + \left(\frac{\partial}{\partial\alpha_2^T}ES_{3i}\right)K_2^{-1}B_{2i}\right].$$

As the estimated ROC curve can be written as functions of $\hat{G}_1$, $\hat{G}_0^{-1}$, and $\hat{\theta}$, the point-wise asymptotic variance of $\widehat{ROC}_x(t)$ is shown in the following theorem:

THEOREM 2: *As the conditions for Theorem 1 and Lemma 1 hold, for any $t \in (0, 1)$,*

$$\sqrt{n} \left(\widehat{ROC}_x(t) - ROC_x(t) \right) \xrightarrow{d} N(0, \Omega_2)$$

where the expression of Ω_2 is given in the Appendix.

The proof of Lemma 1 and Theorem 2 is sketched in the Web Appendix. According to the proof of Theorems 1 and 2, both asymptotic variances consist of two sources of variability: one from the estimating functions S_k $(k = 1, \cdots, 5)$, the other from plugging in the estimated probabilities $\hat{\rho}_i$ and/or $\hat{\pi}_i$. However, a nice property for the DR estimator is that the second source of variability may vanish under some special cases. In other words, the estimated ROC curves with estimated $\hat{\rho}_i$ and $\hat{\pi}_i$ have a similar variance as the estimated ROC curve with the true probabilities ρ_i and π_i. This property is stated in the following Corollary 1, which is proved in the Web Appendix:

COROLLARY 1: *In the DR estimator, if $\hat{\rho}_i$ is estimated with $\sqrt{n}$ consistency, the variances Ω_1 and Ω_2 do not contain the variability of estimating $\hat{\pi}_i$; if $\hat{\pi}_i$ is estimated with $\sqrt{n}$ consistency, Ω_1 and Ω_2 do not contain the variability of estimating $\hat{\rho}_i$.*

4. Simulation studies

We conducted extensive simulation studies and only report the primary results in this section. More results are shown in the Web Appendix. The first simulation compared the performance of the proposed DR, IPW and IB estimators with the estimator in Page and Rotnitzky (2009). The second simulation compares the smooth ROC curve with the empirical estimator in Alonzo and Pepe (2005), when the covariates are not associated with the test result. The

impact of different bandwidth selections is also investigated. Simulation three investigate the model misspecification.

4.1 *Simulation one: comparison to the parametric estimator*

In this simulation, we compare our proposed methods to the doubly robust estimator in Page and Rotnitzky (2009), denoted by PR estimator. Two covariates X_1 and X_2 are generated from $Bernoulli(0.5)$ and $Uniform(-1,1)$, respectively. The true disease status, D, is generated from the conditional distribution, $D \mid X_1, X_2 \sim Bernoulli(\rho)$, where $\text{logit}(\rho) = -0.25 + 0.5X_1 + 0.8X_2$. The test result is generated from $T = \mu(D, X) + \sigma(D, X) \times \epsilon(D)$, where $\mu = 1 + 0.4D + 0.2X_1 + 0.7X_2 + X_1 D + 0.5X_2 D$ and $\sigma = 0.8D + 1.2(1 - D)$. Two scenarios for the residual distribution are simulated: (A) $\epsilon(D) \sim N(0,1)$, (B) $(4.5 + 3\epsilon(0)) \sim \chi^2(4.5)$ and $(8 + 4\epsilon(1)) \sim \chi^2(8)$. Apparently, the test distribution is symmetric in scenario A and skewed in scenario B. The verification indicator is generated from the conditional distribution, $V \mid T, X_1, X_2 \sim Bernoulli(\pi)$, where $\text{logit}(\pi) = -1 + 0.5T + 0.4X_1 + 0.6X_2$. This simulation setup results in about 50% missingness of the gold standard.

[Table 1 about here.]

[Table 2 about here.]

We set the sample size to be 1000. From the data generation, we can see that the disease probability $\rho = \Pr(D = 1 \mid T, X_1, X_2)$ is jointly determined by $D \mid X_1, X_2$ and $T \mid D, X_1, X_2$, i.e., $\rho = \frac{\Pr(D=1|X_1,X_2) f_{T|D=1,X_1,X_2}(t)}{\Pr(D=0|X_1,X_2) f_{T|D=0,X_1,X_2}(t) + \Pr(D=1|X_1,X_2) f_{T|D=1,X_1,X_2}(t)}$. This is a complex function of T, X_1 and X_2, and a linear logistic regression may not estimate the true disease probability well enough. Therefore we also include quadratic terms of T and X_2, as well as pairwise interactions between T, X_1 and X_2. Indeed, under our data generation procedure in scenario A, the quadratic form is exactly the correct model. With χ^2 residual, this disease model still closely approximates the true disease probability. We use the correct verification model

to estimate π. The bandwidth is chosen to be 0.05. We see in simulation study two that the bandwidth has little effect on the performance of our estimators, as long as they are kept small. The results of 500 simulations are reported in Tables 1 and 2. Note that the PR estimator for the location-scale parameters are exactly the same as the DR estimator. We can see from Table 1 that the estimated location and scale parameters always have low bias and close-to-nominal coverage rate. The scale parameters are not as well estimated as the location parameters, which we would expect for most regression analysis.

Table 2 only reports the ROC curve at three 1-specificity levels: 0.1, 0.2 and 0.4 because of two reasons: first, the left end of the ROC curve may be more interesting as those correspond to thresholds with good specificities; second, in our simulation, the ROC curve may be too close to 1 when 1-specificity is greater than 0.4. It is shown that with the normal residual, the bias of the estimated covariate-specific ROC curve is generally small for all the four methods and the coverage rate is close to 95%. However, if the residual distribution is χ^2, the binormal assumption does not hold for the PR estimator. Therefore, the PR estimator is seriously biased and the coverage is much lower than 95%, while the proposed estimators still work well. We notice that if the the true sensitivity is close to 1, the estimated standard errors are not accurate. This is because the data is sparse in estimating the tail probability of the residual distribution.

As for the comparison between the three proposed approaches, the IB estimator has the smallest standard error in general; the DR estimator gains robustness at the cost to efficiency; the IPW estimator is the least efficient among the three, as the estimating functions only use the complete cases. More simulations and discussions on the efficiency issues are given in the Web Appendix. In practice, when the risk factors of disease are well understood, we would recommend the IB estimator; otherwise, the DR estimator is preferable because of its robustness.

We also try smaller sample size ($n = 200$) or lower verification proportion (30%), which suggests that the proposed estimators still perform well with higher than 90% CI coverage rates. These results are similar to Tables 1 and 2, and are omitted here.

4.2 *Simulation two: empirical vs. smooth ROC curve*

When the covariates are not related to the test results, our proposed estimators should be close to that in Alonzo and Pepe (2005), which we refer to as the AP method. The only difference is that we use the kernel smoothing method to estimate the test distribution while they use empirical estimators. In this simulation study, we compare our proposed estimators with the AP method in terms of the mean squared error of the estimated ROC curve. The data generation is similar to scenario A of simulation one, but the mean test result is generated by $\mu = 1 + 0.8D$, which is not affected by the covariates. We only compared the DR estimator here.

[Table 3 about here.]

Table 3 shows that our proposed estimator generally has comparable RMSE as the AP estimator. When the bandwidth is very small, two methods lead to almost identical results. It is also shown that the proposed estimator is not sensitive to the bandwidth selection: for all the bandwidth from 0.01 to 0.2, the bias is small and the RMSE remains at the same magnitude.

4.3 *Simulation three: model misspecification*

In this subsection, we conduct further simulations to examine the model misspecification. First we consider the misspecification of the disease and verification models. Aside from the DR, IPW and IB estimators in simulation one, we consider five additional estimators: DR-V (correct disease model and incorrect verification model), DR-D (correct verification model and incorrect disease model), DR-DV (incorrect disease and verification models),

IPW-V (incorrect verification model), and IB-D (incorrect disease model) estimators. The misspecified verification model ignores the test result, while the misspecified disease model ignores the interactions and quadratic terms. The data generation settings are the same as Scenario A in simulation one.

The simulation results are shown in Figure 1, which plots the ROC curve estimators with covariates $(0,0)$. The results for other covariates levels look similar in general, so we omit the redundant figures. As we expect, the averaging DR, IPW and IB estimators line up with the true ROC curve pretty well. When either the disease model or the verification model is incorrect, DR-D and DR-V estimators are still unbiased, but IB-D and IPW-V estimators both have serious bias. DR-DV estimator is also biased, but the magnitude of bias seems to be a bit smaller. For the CI coverage, all the estimators except DR-DV, IPW-V and IB-D have around 95% coverage rate for the whole range of t. The average standard error is close to the standard deviation of the estimates, suggesting that the asymptotic variance estimator captures the true variability very well.

[Figure 1 about here.]

Furthermore, we check the misspecification of the location-scale model. The data is generated from a transformation model: take $S = 1.6 \times sign(T) \times \sqrt{|T|}$ as the test result instead, where T is the same as in simulation study one. The verification model uses S as a covariate correspondingly. The results for $ROC_{(0,0)}(t)$ are shown in Figure 2. As the location-scale assumption does not hold any more, it is not surprising that all the three proposed estimators are seriously biased.

[Figure 2 about here.]

From the above three sets of simulation studies, we conclude that (1) our proposed methods perform reasonably well in finite sample settings; (2) the PR estimator is seriously biased as the test results do not follow normal distribution; (3) the proposed methods are not sensitive
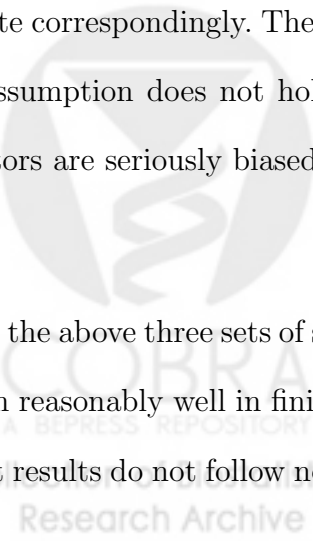
to the bandwidth selection; (4) the DR estimator is more preferable than the IPW and IB estimators as it protects misspecification of either the disease model or the verification model; (5) our proposed method is sensitive to the location-scale model assumption.

5. Example: NACC data

Our proposed method is illustrated using the data collected by National Alzheimer's Coordinating Center (NACC). We included a total of 17,403 deceased patients for our analysis. The test under evaluation is the Mini Mental State Examination (MMSE), which is a brief 30-point questionnaire test used to screen for cognitive impairment. The MMSE score can range from 0 to 30, with lower score indicating more severe impairment. The gold standard ascertainment of AD, based on brain autopsy, is only available for about 31% of the cohort. The missingness may be due to the patients' or their family's decision. We believe that their decision of disease verification may be associated with the demographic characteristics (such as age, gender, race, etc.), but is unlikely to be correlated with their true AD status. So the ignorable missingness assumption seems to be reasonable here. Other covariates extracted from the database are age (continuous variable indicating age at the MMSE test), gender (binary variable with 1 indicating male), race (binary variable with 1 indicating white people), marital status (binary variable with 1 indicating married); clinical diagnosis of AD (binary variable with 1 indicating clinical diagnosed with AD), stroke (binary variable with 1 indicating having stroke before), Parkinson's disease (binary variable with 1 indicating presence of the disease), and depression (binary variable with 1 indicating presence of the disease). Figure 3 displays the distribution of the MMSE score for all the patients and for stratified verification and disease status. The distribution of the test result seems to be irregular, so it is hard to assume any parametric distribution. The test distributions for the cases and controls are very different too.

[Figure 3 about here.]

We transformed the MMSE score using $(30 - T)/5$ so that a diseased subject tends to have larger test result, and transform the age using $(age - 70)/10$ so that the reported coefficients are in the appropriate magnitude. All the aforementioned covariates, as well as the MMSE score are included in the verification model. The disease model also includes the quadratic term of the test score, as well as the interaction between the test and the covariates. For modelling the MMSE score, the location model has the main effects of D and X, as well as their interactions, while the scale model only has the main effect of D with log link. The estimated coefficients of the location and scale model are given in Table 4 using DR, IPW and IB approaches. The results from the three methods generally coincide with each other. The DR estimator identifies main effects of race, clinical AD, and true disease status to be significant, indicating that these variables affects the magnitude of the test score. The $race \times D$ interaction is significant, while gender, clinical AD and depression have marginally insignificant interactions with D.

[Table 4 about here.]

We take the bandwidth to be 0.02 in estimating the ROC curve. Hence the covariate-specific ROC curve could be plotted for every covariate level. For example, Figure 4 shows the DR estimates of the two specific ROC curves: one for 70 years old non-white female with other covariates being 0, the other for 60 years old white male with depression and other covariates being 0. The 95% CI's are also plotted. The results show that the classification ability of the MMSE test could be very different according to the covariates stratification. Although the test result can only take integer values from 0 to 30, the estimated ROC curve is smooth as the kernel estimating equations are used for the distributions. The empirical version of the ROC curve is more coarse, which could have at most 31 jumps.

[Figure 4 about here.]

In Figure 5, we plot the DR and PR estimators of the area under ROC curve (AUC) as a function of age, with other binary covariates fixed. The solid line is for a white married male with clinical AD diagnosis and no other diseases; the dashed line is for a white not-married female with clinical AD diagnosis and no other diseases. These two covariates groups are the most prevalent in the NACC data set. We can see that the AUC is increasing by age with almost linear trend. We also find some discrepancy between DR and PR estimators, especially for older patients.

[Figure 5 about here.]

In screening for dementia or mild cognitive impairment (MCI), previous literatures suggest that the AUC for MMSE score is usually above 0.7 (Kim, *et al.*, 2005; Isella, *et al.*, 2006; McDowell, *et al.*, 1997). However, we found that the MMSE score is not as promising in detecting AD, especially in younger patients. This motivates for further study of new biomarkers or combined biomarkers to improve early diagnosis of Alzheimer's disease. On the other hand, given the simplicity and low cost of the MMSE test, it is still of great use in screening for cognitive impairment in practice.

6. Discussion

In this paper, we have proposed to estimate the covariate-specific ROC curve semiparametrically when the gold standard is subject to missingness. The form of the ROC curve is flexible as both the link and the baseline functions are unknown and estimated from data. Three approaches are proposed to adjust for the verification bias: DR, IPW and IB estimators, which use different weights in the estimating equations. The disease probability and the verification probability are the key components in constructing the weighted estimating equations. The DR estimator only requires that either disease or verification model is correctly specified to have a consistent ROC curve estimator. The doubly robust property allows two shots to
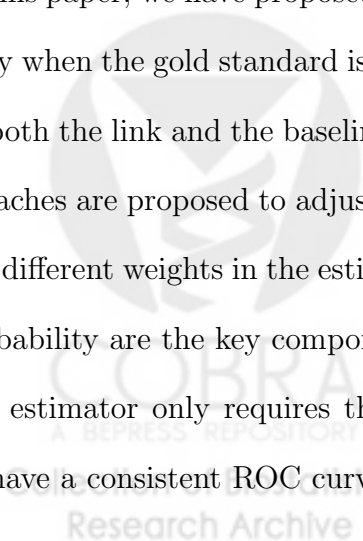
assume the correct model, and is favorable in practice. Otherwise, if the disease probability could be modelled correctly, the IB estimator is the most efficient. The IPW estimator is the least efficient among the three estimators. Our estimating procedures are based on the location-scale model framework, where the link and baseline functions of the ROC curve are just a distribution function and a quantile function. The location-scale parameters as well as the unspecified residual distributions are estimated, which determine the estimated ROC curve together.

Although we are focusing on the MAR verification process in this paper, the extension to nonignorable missingness (NI) is straightforward. Under the NI assumption, the observed data likelihood involves both π_i and ρ_i, that usually need to be estimated together. We can adopt the likelihood-based estimation of the nonignorable selection model in Liu and Zhou (2010). Or as in Rotnitzky *et al.* (2006) and Fluss *et al.* (2009), we specify the odds ratio of verification given disease, and then estimate π_i and ρ_i separately. With the estimated disease and verification probabilities, our proposed weighted estimating functions still work with slight modifications: for DR and IB approach, we replace ρ_i with $\rho_{i0} \equiv \Pr(D_i = 1 | V_i = 0, T_i, X_i)$. The resultant asymptotic variances take the similar form as the MAR case.

An alternative method might be a transformation model, i.e., we assume that $h(T)$ follows the distribution F (usually specified up to some parameters), where the smooth transformation h is left unspecified. The direct ROC curve estimation method in Cai and Pepe (2002) is in fact a special case of the transformation model. They assume that h is the distribution function of the test result for the controls. The direct estimation has the advantage of easy interpretation of model parameters, i.e., the effects of covariates on the ROC curve. The indirect estimation is relatively easy in modelling the location and scale of the test result. Pepe (1998) pointed out that the indirect estimation yields more efficient estimators than the direct estimation. Indeed, our proposed estimating equations for the location-scale

parameters are just the gaussian score equations. Therefore, when the test distribution is close to normal, the location-scale model would be the most efficient. For highly skew data, the performance of direct and indirect ROC curve estimation is worthwhile for further exploration.

It follows from Nadaraya (1964) that the bias of the kernel CDF estimator is negligible relative to its variance, as long as the bandwidth is kept small enough. This property guarantees the $\sqrt{n}$ consistency of the estimated case and control distributions, and hence the $\sqrt{n}$ consistency of the estimated ROC curve. Another advantage of kernel smoothing is that it gives smooth ROC curve estimates as often desired.

7. Supplementary Material

Web Appendices referenced in Sections 3 and 4 are available under the Paper Information link at the *Biometrics* website `http://www.biometrics.tibs.org`.
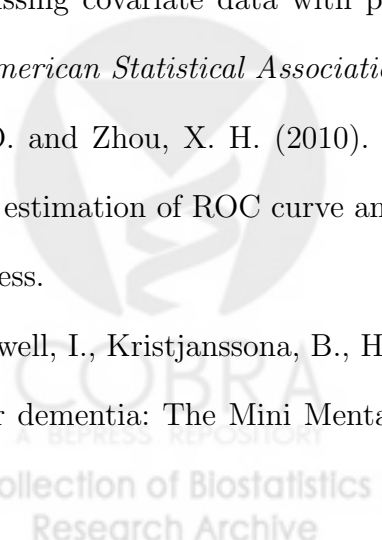
References

Alonzo, T. A. and Pepe, M. S. (2005). Assessing accuracy of a continuous screening test in the presence of verification bias. *Jour of the Royal Statistical Society - Series C (Applied Statistics)*, **54**, 173-190.
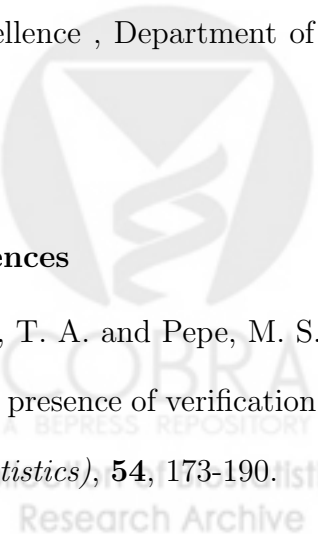
Begg, C. B. and Greenes, R. A. (1983). Assessment of diagnostic tests when disease verification is subject to verification bias. *Biometrics*, **39**, 207-215.

Cai, T. and Pepe, M. S. (2002). Semiparametric receiver operating characteristic analysis to evaluate biomarkers for disease. *Journal of the American Statistical Association*, **97**, 1099-1107.

Fluss, R., Reiser, B., Faraggi, D. and Rotnitzky, A. (2009). Estimation of the ROC curve under verification bias. *Biometrical Journal*, **51**, 475-490.

Gary, R., Begg, C. B. and Greenes, R. A. (1983). Construction of receiver operating characteristic curves when disease verification is subject to selection bias. *Medical Decision Making*, **4**, 151-164.

Isella, V., Villa, L., Russo, A., Regazzoni, R., Ferrarese, C. and Appollonio, I. M. (2006). Discriminative and predictive power of an informant report in mild cognitive impairment. *Journal of Neurology, Neurosurgery & Psychiatry*, **77**, 166-171.

Kim, K. W., Lee, D. Y., Jhoo, J. H., Youn, J. C., Suh, Y. J., Jun, Y. H., Seo, E. H. and Woo, J. I. (2005). Diagnostic accuracy of mini-mental status examination and revised hasegawa dementia scale for Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, **19**, 324-330.

Lipsitz, S. R., Ibrahim, J. G. and Zhao, L. P. (1999). A weighted estimating equation for missing covariate data with properties similar to maximum likelihood. *Journal of the American Statistical Association*, **94**, 1147-1160.

Liu, D. and Zhou, X. H. (2010). A model for adjusting for nonignorable verification bias in estimation of ROC curve and its area with likelihood-based approach. *Biometrics*, In press.

McDowell, I., Kristjanssona, B., Hilla, G. B. and Hebertb, R. (1997). Community screening for dementia: The Mini Mental State Exam (MMSE) and modified Mini-Mental State

Exam (3MS) compared. *CJournal of Clinical Epidemiology*, **50**, 377-383.

Nadaraya, E. A. (1964). Some new estimates for distribution functions. *Theory of Probability and its Applications*, **9**, 497–500.

Page, J. H. and Rotnitzky, A. (2009). Estimation of the disease-specific diagnostic marker distribution under verification bias. *Computational Statistics and Data Analysis*, **53**, 707-717.

Pepe, M. S. (1998). Three approaches to regression analysis of receiver operating characteristic curves for continuous test results. *Biometrics*, **54**, 124-135.

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

Punglia, R. S., D'Amico, A. V., Catalona, W. J., Roehl, K. A., and Kuntz, K. M. (2003). Effect of verification bias on screening for prostate cancer by measurement of prostate-specific antigen. *New England Journal of Medicine*, **349**, 335-342.

Rodenberg, C. A. and Zhou X. H. (2000). ROC curve estimation when covariates affect the verification process. *Biometrics*, **56**,1256-1262.

Rotnitzky, A., Faraggi, D. and Schisterman, E. (2006). Doubly robust estimation of the area under the receiver-operating characteristic curve in the presence of verification bias. *Journal of the American Statistical Association*, **101**, 1276-1288.

Zheng, Y., Barlow, W. E., and Cutter, G. (2005). Assessing accuracy of mannography in the presence of verification bias and intrareader correlation. *Biometrics*, **61**, 259-268.

Zhou, X. H. (1996). A nonparametric ML estimate of an ROC curve area corrected for verification bias. *Biometrics*, **52**, 310-316.

Zhou, X. H. (1998). Comparing correlated areas under the ROC curves of two diagnostic tests in the presence of verification bias. *Biometrics*, **54**, 453-470.

Zhou, X. H., Obuchowski, N. A. and McClish, D. K. (2002). *Statistical Methods in Diagnostic*

Medicine. New York: Wiley.

Zhou, X. H. and Rodenberg, C. A. (1998). Estimating an ROC curve in the presence of non-ignorable verification bias. *Communications in Statistics*, **27**, 635-657.

Figure 1. The true and estimated ROC curve, empirical standard deviation, average standard error, and the CI coverage rate for the estimated ROC curves using DR, DR-D, DR-V, DR-DV, IPW, IPW-V, IB and IB-D estimators.
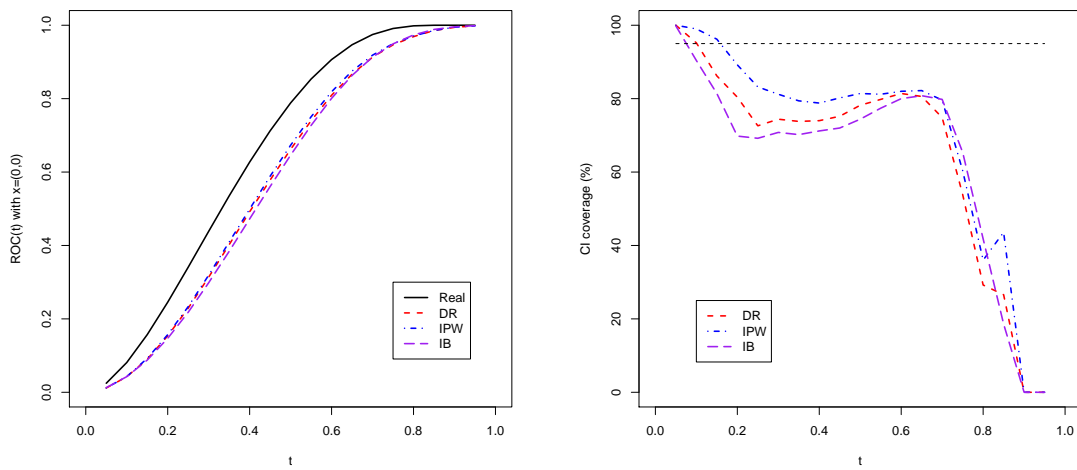
Figure 2. The true and estimated ROC curve and the CI coverage rate for the DR, IPW and IB estimators under transformation model.
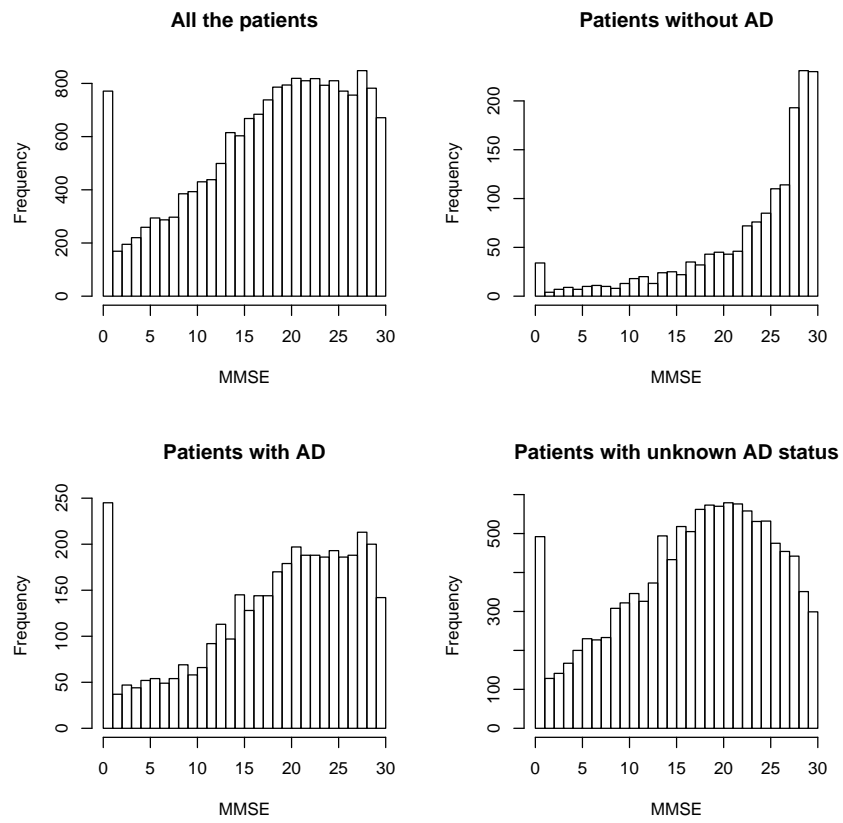
Figure 3. MMSE distribution by AD status and verification.
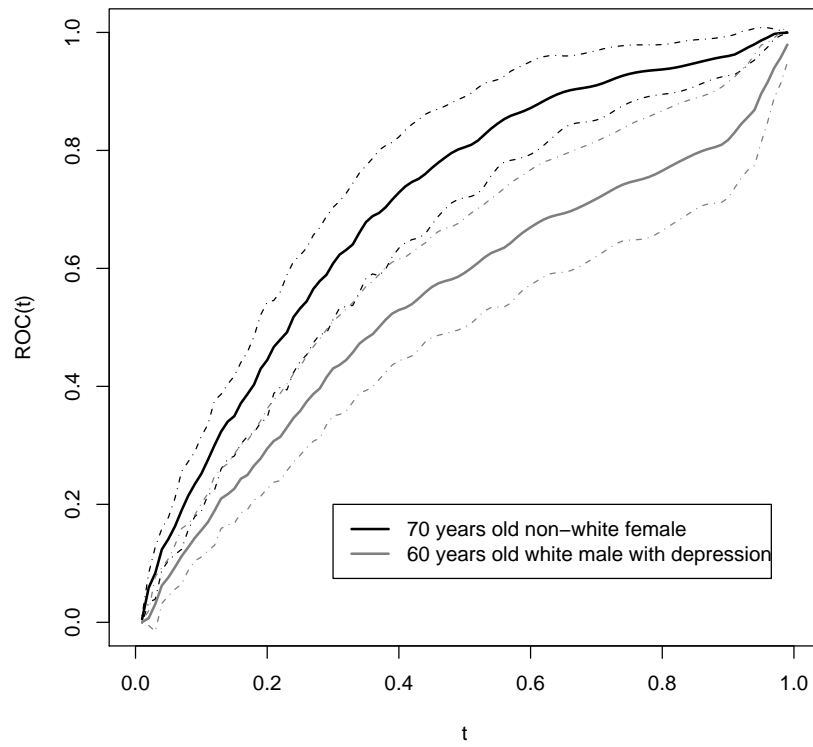
Figure 4. The covariate-specific ROC curves.
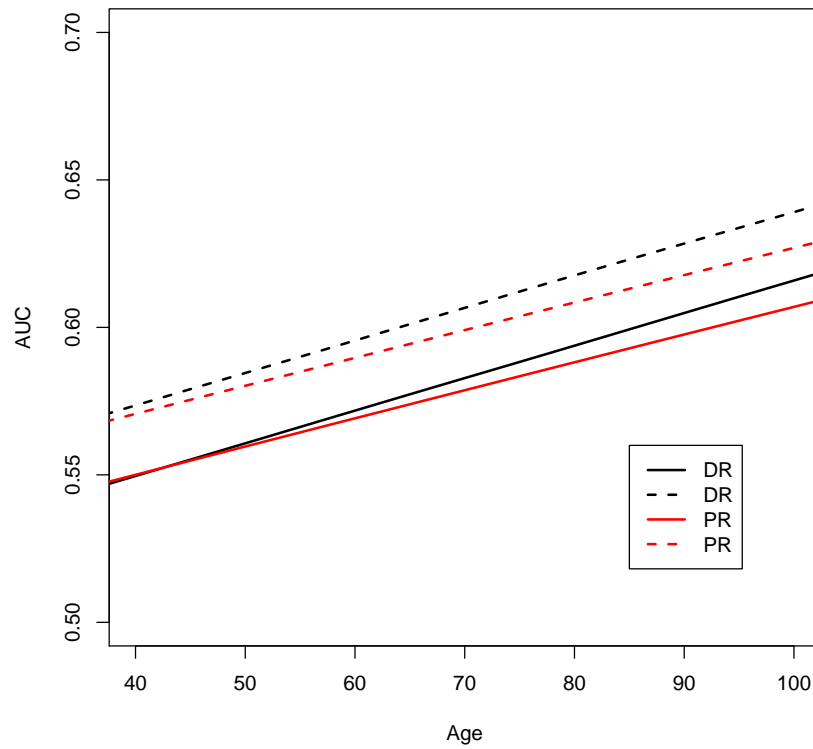
Figure 5. DR and PR estimators for the covariate-specific AUC as a function of age: the solid line is for a white married male with clinical AD diagnosis and no other diseases; the dashed line is for a white not-married female with clinical AD diagnosis and no other diseases.

Table 1

The bias (in percentage to the true value), empirical standard deviation (SD), average standard error (SE), and the CI coverage rate for the location and scale parameters.

					Location				Scale	
			Intercept	D	X_1	X_2	$X_1 \times D$	$X_2 \times D$	Intercept	D
Normal	Bias (%)	DR	-0.5	2.3	1.9	-0.8	-0.8	-0.1	-3.0	-1.2
		IPW	0.1	0.3	4.6	-0.8	-1.1	-0.4	-10.5	1.8
		IB	-0.4	1.6	1.9	-0.9	-0.7	0.6	-3.2	-0.3
	SD	DR	0.091	0.151	0.130	0.112	0.187	0.164	0.035	0.064
		IPW	0.132	0.176	0.206	0.191	0.242	0.216	0.067	0.085
		IB	0.086	0.138	0.124	0.105	0.172	0.145	0.035	0.059
	SE	DR	0.085	0.138	0.123	0.109	0.174	0.157	0.035	0.058
		IPW	0.126	0.163	0.196	0.171	0.225	0.199	0.057	0.074
		IB	0.083	0.133	0.120	0.105	0.165	0.147	0.034	0.055
	Coverage (%)	DR	95.0	92.2	94.2	94.0	93.6	93.6	93.6	93.0
		IPW	94.2	93.2	94.0	91.4	94.2	93.2	87.4	91.4
		IB	95.6	93.6	94.6	93.6	95.2	94.2	93.2	93.6
χ^2	Bias (%)	DR	0.3	-1.6	-7.2	-1.2	2.1	0.9	-2.4	-1.5
		IPW	1.1	-3.5	-10.9	-1.6	2.8	1.1	-4.2	-0.7
		IB	0.1	-0.3	-0.2	-0.1	-0.8	-1.0	-5.2	5.6
	SD	DR	0.083	0.130	0.117	0.105	0.157	0.149	0.052	0.072
		IPW	0.114	0.150	0.169	0.148	0.195	0.177	0.055	0.074
		IB	0.082	0.128	0.119	0.105	0.161	0.146	0.052	0.074
	SE	DR	0.082	0.126	0.120	0.106	0.162	0.147	0.050	0.070
		IPW	0.105	0.140	0.164	0.145	0.192	0.172	0.055	0.074
		IB	0.081	0.124	0.119	0.105	0.162	0.144	0.050	0.071
	Coverage (%)	DR	95.0	94.4	96.0	95.6	95.0	94.2	93.4	94.4
		IPW	92.4	93.4	94.2	93.4	94.6	94.4	95.0	94.6
		IB	94.2	94.8	95.6	95.2	95.2	94.6	93.0	91.8

Table 2

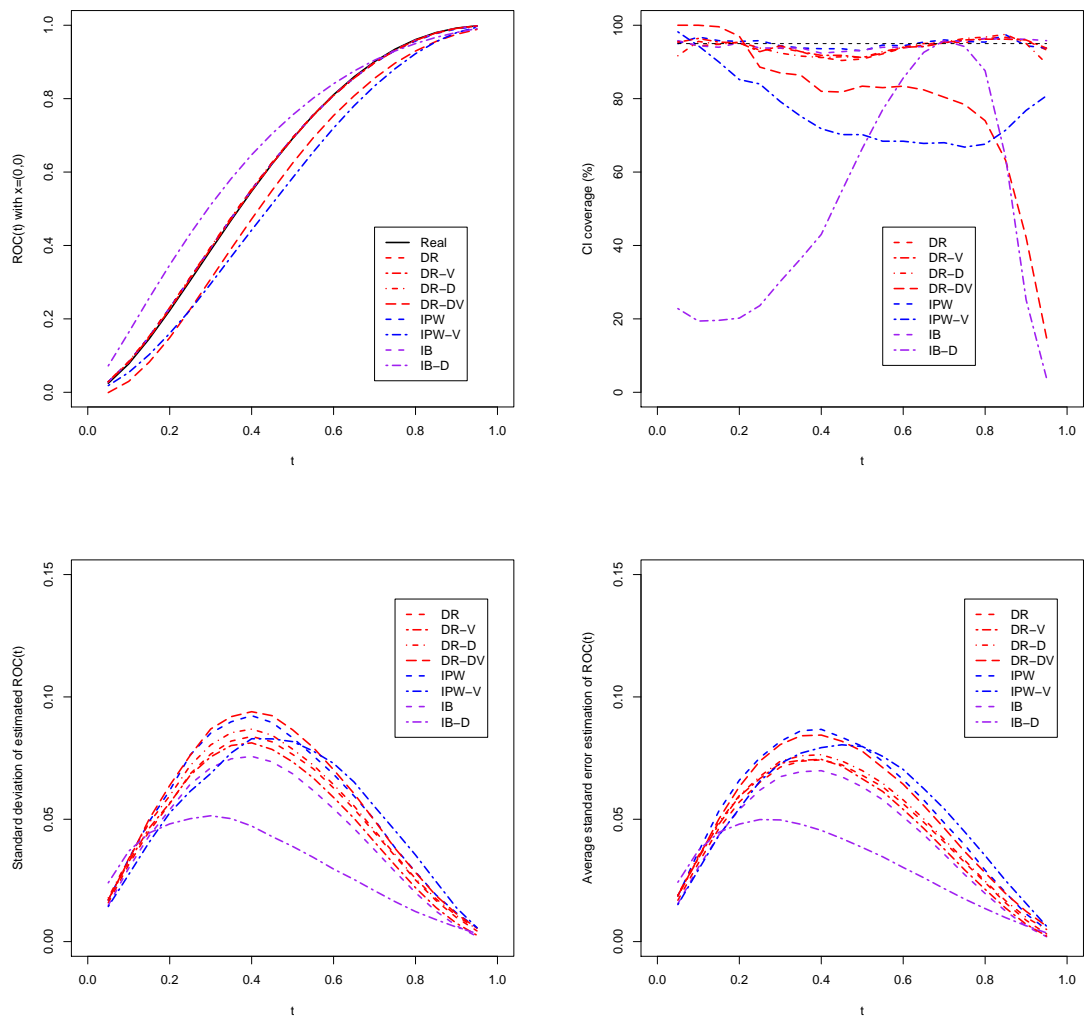*The bias (in percentage to the true value), empirical standard deviation (SD), average standard error (SE), and the CI coverage rate for the estimated ROC curves by DR, IPW and IB estimators.*

|		$ROC_x(0.1)$				$ROC_x(0.2)$				$ROC_x(0.4)$			
		$x=(0,0)$	$(0,0.5)$	$(1,0)$	$(1,0.5)$	$(0,0)$	$(0,0.5)$	$(1,0)$	$(1,0.5)$	$(0,0)$	$(0,0.5)$	$(1,0)$	$(1,0.5)$
Normal	Bias (%)												
	DR	6.8	4.4	0.0	-0.1	2.8	2.0	0.3	-0.0	0.9	0.6	-0.1	-0.1
	IPW	6.7	4.9	-0.3	-0.5	3.2	2.2	0.0	-0.5	0.8	0.4	-0.2	-0.3
	IB	6.4	4.6	0.4	0.3	2.6	2.2	0.3	0.1	0.6	0.5	-0.1	-0.1
	PR	6.1	4.6	1.0	0.6	3.2	2.2	0.4	0.1	1.1	0.6	0.0	-0.1
	SD												
	DR	0.031	0.044	0.073	0.082	0.056	0.068	0.059	0.055	0.084	0.073	0.030	0.021
	IPW	0.034	0.053	0.091	0.107	0.062	0.082	0.078	0.078	0.092	0.087	0.037	0.030
	IB	0.030	0.043	0.069	0.075	0.054	0.066	0.056	0.050	0.076	0.070	0.027	0.019
	PR	0.029	0.042	0.063	0.072	0.056	0.065	0.055	0.050	0.075	0.066	0.026	0.017
	SE												
	DR	0.032	0.046	0.069	0.078	0.057	0.070	0.060	0.056	0.074	0.070	0.031	0.023
	IPW	0.037	0.056	0.087	0.099	0.066	0.084	0.075	0.072	0.087	0.085	0.039	0.030
	IB	0.030	0.044	0.065	0.072	0.054	0.066	0.055	0.051	0.070	0.066	0.028	0.020
	PR	0.030	0.044	0.058	0.068	0.056	0.067	0.049	0.048	0.068	0.063	0.022	0.017
	Coverage (%)												
	DR	94.4	95.4	93.4	94.2	95.2	94.4	95.4	95.6	91.2	93.2	95.6	97.6
	IPW	96.8	96.4	93.2	94.6	95.6	95.2	94.6	95.2	93.6	93.8	96.6	96.8
	IB	94.4	94.6	93.6	94.6	95.2	93.8	94.0	96.0	92.4	93.6	94.8	97.4
	PR	95.2	95.0	92.6	93.8	95.0	94.4	92.0	95.0	92.8	93.6	91.4	94.6
χ^2	Bias (%)												
	DR	3.9	3.4	4.3	4.2	1.3	1.3	1.4	0.5	-0.1	-0.5	-0.3	-0.0
	IPW	4.1	4.0	4.7	4.7	1.2	1.7	1.7	0.6	-0.2	-0.6	-0.3	-0.1
	IB	10.2	10.1	7.3	6.1	8.5	7.4	2.0	-0.2	1.2	-0.6	-1.9	-0.9
	PR	-0.7	14.5	29.7	24.2	-9.0	-3.7	-4.5	-7.0	-13.2	-13.1	-7.9	-4.5
	SD												
	DR	0.030	0.041	0.074	0.092	0.060	0.074	0.077	0.066	0.082	0.077	0.009	0.003
	IPW	0.031	0.044	0.080	0.102	0.066	0.083	0.086	0.075	0.092	0.088	0.009	0.003

Table 3

The comparison of our proposed DR estimator and the AP estimator in terms of bias (in percentage of the true value), 100 times empirical standard deviation (SD), and 100 times root mean squared error (RMSE).

		$Bias(\%)$		SD		$RMSE$	
		$ROC(0.1)$	$ROC(0.2)$	$ROC(0.1)$	$ROC(0.2)$	$ROC(0.1)$	$ROC(0.2)$
DR	$h = 0.01$	0.4	-0.1	5.0	4.4	5.0	4.4
	$h = 0.05$	0.4	-0.1	4.9	4.3	4.9	4.3
	$h = 0.1$	-0.1	-0.2	4.8	4.2	4.8	4.2
	$h = 0.2$	-1.5	-0.9	4.5	4.0	4.6	4.1
AP		0.1	-0.2	5.0	4.4	5.0	4.4

Table 4

The estimated location and scale parameters with the associated standard errors for the NACC data.

	DR	IPW	IB
Intercept	**1.394 (0.125)**	**1.452 (0.199)**	**1.408 (0.131)**
Age	-0.064 (0.039)	**-0.114 (0.051)**	-0.054 (0.032)
Gender	-0.049 (0.079)	0.046 (0.096)	-0.028 (0.074)
Race	**-0.303 (0.131)**	**-0.525 (0.203)**	**-0.304 (0.128)**
Marital status	0.083 (0.086)	**0.219 (0.102)**	0.031 (0.076)
Clinical AD	**1.120 (0.077)**	**1.160 (0.097)**	**1.128 (0.070)**
Stroke	0.129 (0.095)	0.100 (0.125)	0.129 (0.083)
Parkinson's	0.186 (0.138)	0.228 (0.153)	**0.292 (0.124)**
Depression	-0.007 (0.108)	0.144 (0.156)	-0.006 (0.083)
D	**0.984 (0.167)**	**1.090 (0.250)**	**0.982 (0.170)**
$D \times Age$	0.049 (0.052)	0.085 (0.063)	0.035 (0.041)
$D \times Gender$	-0.146 (0.099)	-0.065 (0.119)	**-0.180 (0.090)**
$D \times Race$	**-0.322 (0.158)**	-0.339 (0.240)	**-0.330 (0.152)**
$D \times Marital\ status$	0.039 (0.108)	-0.090 (0.127)	0.110 (0.094)
$D \times Clinical\ AD$	-0.150 (0.106)	-0.154 (0.127)	-0.166 (0.095)
$D \times Stroke$	-0.124 (0.120)	-0.082 (0.156)	-0.115 (0.105)
$D \times Parkinson's$	0.074 (0.172)	0.012 (0.190)	-0.058 (0.155)
$D \times Depression$	-0.225 (0.135)	-0.119 (0.180)	-0.153 (0.102)
Intercept	**0.301 (0.024)**	**0.391 (0.028)**	**0.310 (0.022)**
D	**0.120 (0.027)**	**0.149 (0.031)**	**0.109 (0.025)**