Electronic Journal of Statistics
Vol. 10 (2016) 3063–3113
ISSN: 1935-7524
DOI: 10.1214/16-EJS1202

Bias–corrected methods for estimating the receiver operating characteristic surface of continuous diagnostic tests

Khanh To Duc, Monica Chiogna and Gianfranco Adimari

Department of Statistical Sciences, University of Padua
Via C. Battisti, 241-243, 35121 Padua, Italy
e-mail: toduc@stat.unipd.it; monica.chiogna@unipd.it; gianfranco.adimari@unipd.it

Abstract: Verification bias is a well-known problem that may occur in the evaluation of predictive ability of diagnostic tests. When a binary disease status is considered, various solutions can be found in the literature to correct inference based on usual measures of test accuracy, such as the receiver operating characteristic (ROC) curve or the area underneath. Evaluation of the predictive ability of continuous diagnostic tests in the presence of verification bias for an ordinal three-class disease status is here discussed. In particular, several verification bias-corrected estimators of the ROC surface and of the volume underneath are proposed. Consistency and asymptotic normality of the proposed estimators are established and their finite sample behavior is investigated by means of Monte Carlo simulation studies. Two illustrations are also given.

MSC 2010 subject classifications: Primary 62C99; secondary 62P10.
Keywords and phrases: Verification bias, missing at random, ROC surface analysis, true class fractions.

Received June 2016.

Contents

1. Introduction

Before applying a diagnostic test in clinical settings, a rigorous statistical assessment of its performance in discriminating the disease status from the non-disease status is necessary. For a continuous-scale test T, the diagnosis is dependent upon whether the test result is above or below a specified cut point c. Assuming, without loss of generality, that higher test values indicate a higher likelihood of disease, a result is called positive if its value exceeds the cut point, and negative otherwise. A positive test indicates presence of the disease.

At a fixed cut point c, the accuracy of the test can be evaluated by its true positive rate (TPR) and its true negative rate (TNR), which are defined as the probabilities that the test correctly identifies the diseaded and non-diseaded subjects, respectively. The Receiver Operating Characteristic (ROC) curve is the plot of TPR versus 1-TNR by varying the cut point c. The ROC curve is monotone and, provided that a correct specification of class ordering is done prior to the decision making rule, lies in the upper triangle of the unit square, which consists of three vertices $(0,0), (0,1)$ and $(1,1)$. The shape of ROC curve allows to evaluate the ability of the test. For example, a ROC curve equal to a straight line joining points $(0,0)$ and $(1,1)$ represents a diagnostic tests which is the random guess. A commonly used summary measure that aggregates performance information across the range of possible cut points is the area under ROC curve (AUC). Under correct ordering, values of AUC range from 0.5, suggesting that the test is no better than chance alone, to 1, indicating a perfect test.

Clearly, the ROC curve and the AUC of a test under assessment are unknown and the statistical evaluation of the test requires suitable inferential procedures. See, for example, Zhou, McClish and Obuchowski [18] and Pepe [14] as general references. In principle, knowledge of the true disease status of the subjects under study is obtained by the most accurate available test, called gold standard (GS) test. In practice, there may be many drawbacks to implement the GS test, which can be too expensive, or too invasive, or both for regular use. Thus, often only a subset of patients undergoes disease verification and the decision to send a patient to verification typically depends on the test result and other patient characteristic. Statistical evaluations based only on data from subjects with

verified disease status are typically biased. This bias is known as verification bias.

In the last fifteen years, various methods have been developed to deal with the verification bias problem, most of which assume that the true disease status, if missing, is missing at random (MAR, Little and Rubin [11]). Among the others, we cite the papers by Adimari and Chiogna [1], Alonzo and Pepe [3], Fluss et al. [6], Gu, Ghosal and Kleiner [7] and Rotnitzky, Faraggi and Schisterman [15]. In particular, Alonzo and Pepe [3] proposed four types of partially parametric estimators of TPR and TNR, i.e., full imputation (FI), mean score imputation (MSI), inverse probability weighting (IPW) and semiparametric efficient (SPE) estimators.

In some medical studies, however, the disease status often involves more than two categories; for example, Alzheimer's dementia can be classified into three categories (see Chi and Zhou [4] for more details). In such situations, quantities used to evaluate the accuracy of a diagnostic test are the true class fractions (TCF's). These are well defined as a generalization of TPR and TNR. In a three–class diagnostic problem, given a pair of cut points (c_1, c_2), with $c_1 < c_2$, subjects are classified into class 1 if $T < c_1$; class 2 if $c_1 \leq T < c_2$; and class 3 otherwise. This implies that the disease classes are ordered with respect to the test result, a condition often referred to as monotone ordering. The true class fractions of the test T at (c_1, c_2) are defined as

$$\begin{aligned} \mathrm{TCF}_1(c_1) &= \Pr(T < c_1 | \text{class } 1) = 1 - \Pr(T \geq c_1 | \text{class } 1), \\ \mathrm{TCF}_2(c_1, c_2) &= \Pr(c_1 < T < c_2 | \text{class } 2) \\ &= \Pr(T \geq c_1 | \text{class } 2) - \Pr(T \geq c_2 | \text{class } 2), \\ \mathrm{TCF}_3(c_2) &= \Pr(T > c_2 | \text{class } 3) = \Pr(T \geq c_2 | \text{class } 3). \end{aligned}$$

The plot of (TCF_1, TCF_2, TCF_3) by varying the pair (c_1, c_2) produces the ROC surface of T in the unit cube. Scurfield [16] and Nakas and Yiannoutsos [13] mentioned that a ROC surface is well defined as a generalization of the ROC curve. Indeed, the projection of ROC surface to the plane defined by TCF_2 versus TCF_1 yields the ROC curve between classes 1 and 2. Similarly, on projecting the ROC surface to the plane defined by the axes TCF_2 and TCF_3, the ROC curve between classes 2 and 3 is produced (see also [12]). The ROC surface will be the triangular plane with vertices $(0, 0, 1), (0, 1, 0)$, and $(1, 0, 0)$ if all of three TCF's are equal for every pair (c_1, c_2). In this case, we say that the diagnostic test is the random guess, again. In practice, under a correct specification of class ordering, one can imagine that the graph of ROC surface lies in the unit cube and above the plane of the triangle with three vertices $(0, 0, 1), (0, 1, 0)$, and $(1, 0, 0)$. A summary of the overall diagnostic accuracy of the test under consideration is the volume under the ROC surface (VUS) which can be seen as a generalization of the AUC. Under correct ordering, values of VUS vary from 1/6 to 1, ranging from chance to perfect diagnostic tests.

Nakas and Yiannoutsos [13] and Nakas [12] gave some interesting results about ROC surface analysis in absence of verification bias. In particular, the authors formularized the ROC surface by a functional form and proposed a

nonparametric approach for VUS estimation. Again without verification bias, parametric estimation of VUS is supplied in the work of Xiong et al. [17], where the assumption of normality distribution was used, whereas Li and Zhou [9] tackled the nonparametric and semi-parametric estimation of the ROC surface. Li, Zhou and Fine [10] proposed a regression approach to ROC surface, and in Kang and Tian [8] a kernel smoothing based approach for estimation of VUS is employed.

The issue of correcting for the verification bias in ROC surface analysis is very scarcely considered in the statistical literature. Until now, only Chi and Zhou [4] discussed about the issue. The authors proposed maximum likelihood estimates for ROC surface and VUS. However, these results only concern ordinal diagnostic tests. This motivated us to develop bias-corrected methods for continuous diagnostic tests with three–class disease status.

In this paper, we propose several verification bias-corrected estimators of TCF_1, TCF_2 and TCF_3 for continuous diagnostic tests. The proposed estimators are the extension of FI, MSI, IPW and SPE estimators for the ROC curves in Alonzo and Pepe [3]. The new estimators allow to obtain bias-corrected ROC surfaces. Corresponding estimators of the VUS are also presented. Consistency and asymptotic normality of the proposed estimators are established under the MAR assumption.

The rest of paper is organized as follows. In Section 2, we review the estimators of ROC curves discussed in Alonzo and Pepe [3]. The proposed extension, giving bias-corrected estimators of the ROC surface and of VUS, is presented in Section 3, along with the relevant asymptotic results. In Section 4, some simulation results are produced and two applications of the methods are contained in Section 5. Finally, conclusions are drawn in Section 6.

2. Background

In this section, we review the approaches presented in Alonzo and Pepe [3] for bias–corrected ROC analysis in two–class problems.

2.1. Notation and assumption

Let us consider a study with n subjects, for whom the result of a continuous test T is available. The patient's true condition (or disease status), D, is defined by a GS test. D is a binary variable, that is 0 if the subject is healthy and 1 in case of disease. Further, let V be a binary verification status of a patient, such that $V = 1$ if he/she is underwent the GS test, and $V = 0$ otherwise. In practice, some information, other than the test results, can be obtained for each patient. Let A be a covariate vector for a patient, that may be associated with both D and V. For the sake of reader's convenience, without loss of generality, in what follows we will consider A to be univariate. We assume that the verification status V and the response D are mutually independent given the test result T and covariate A, i.e., $\Pr(V|T, A) = \Pr(V|D, T, A)$ or, equivalently, $\Pr(D|T, A) = \Pr(D|V, T, A)$. This assumption corresponds to the MAR assumption.

2.2. Bias correction for ROC curve

Let FPR = 1-TNR. When all subjects are verified by GS, we have a full (or complete) data set. For a given cut point c, TPR and FPR are

$$\begin{aligned} \mathrm{TPR}(c) &= \Pr(T \geq c | D = 1) = \frac{\Pr(T \geq c, D = 1)}{\Pr(D = 1)} = \frac{\beta_1}{\theta}, \\ \mathrm{FPR}(c) &= \Pr(T \geq c | D = 0) = \frac{\Pr(T \geq c, D = 0)}{\Pr(D = 0)} = \frac{\beta_0}{1 - \theta}, \end{aligned} \tag{2.1}$$

in which we define $\beta_0 = \Pr(T \geq c, D = 0)$, $\beta_1 = \Pr(T \geq c, D = 1)$ and $\theta = \Pr(D = 1)$. Then, one can employ the empirical estimators $\hat{\beta}_0, \hat{\beta}_1$ and $\hat{\theta}$ to obtain the nonparametric estimators of TPR and FPR

$$\widehat{\mathrm{TPR}}(c) = \frac{\hat{\beta}_1}{\hat{\theta}} = \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c) D_i}{\sum_{i=1}^{n} D_i}, \quad \widehat{\mathrm{FPR}}(c) = \frac{\hat{\beta}_0}{1 - \hat{\theta}} = \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c)(1 - D_i)}{\sum_{i=1}^{n} (1 - D_i)}, \tag{2.2}$$

where $\mathrm{I}(\cdot)$ is the indicator function.

If not all patients have their disease status verified, the nonparametric estimators (2.2) can not be computed. If one computes the Naïve estimators, i.e., estimators (2.2) based only on verified subjects, typically gets estimates that are biased and inconsistent.

Alonzo and Pepe [3] proposed four partially parametric estimators to assess continuous diagnostic (or screening) tests under the MAR assumption. In particular, FI estimators of $\mathrm{TPR}(c)$ and $\mathrm{FPR}(c)$ are

$$\widehat{\mathrm{TPR}}_{\mathrm{FI}}(c) = \frac{\hat{\beta}_{1,\mathrm{FI}}}{\hat{\theta}_{\mathrm{FI}}} = \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c) \hat{\rho}_i}{\sum_{i=1}^{n} \hat{\rho}_i},$$

$$\widehat{\mathrm{FPR}}_{\mathrm{FI}}(c) = \frac{\hat{\beta}_{0,\mathrm{FI}}}{1 - \hat{\theta}_{\mathrm{FI}}} = \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c)(1 - \hat{\rho}_i)}{\sum_{i=1}^{n} (1 - \hat{\rho}_i)}.$$

Here, the estimates $\hat{\rho}_i$ of $\rho_i = \Pr(D_i = 1 | T_i, A_i)$ are obtained by using some suitable parametric model (e.g., logistic regression model) computed from verified subjects. MSI estimators only imputes the disease status for subjects who did not undergo the GS, resulting to be

$$\widehat{\mathrm{TPR}}_{\mathrm{MSI}}(c) = \frac{\hat{\beta}_{1,\mathrm{MSI}}}{\hat{\theta}_{\mathrm{MSI}}} = \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c) \{V_i D_i + (1 - V_i) \hat{\rho}_i\}}{\sum_{i=1}^{n} \{V_i D_i + (1 - V_i) \hat{\rho}_i\}}, \tag{2.3}$$

$$\widehat{\mathrm{FPR}}_{\mathrm{MSI}}(c) = \frac{\hat{\beta}_{0,\mathrm{MSI}}}{1-\hat{\theta}_{\mathrm{MSI}}} = \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c)\left\{V_i(1-D_i)+(1-V_i)(1-\hat{\rho}_i)\right\}}{\sum_{i=1}^{n}\left\{V_i(1-D_i)+(1-V_i)(1-\hat{\rho}_i)\right\}}.$$

IPW method weights each verified subject by the inverse of the conditional verification probability $\pi_i = \Pr(V_i = 1|T_i, A_i)$ (i.e. the probability that the subject is selected for verification). Therefore, the estimators are

$$\widehat{\mathrm{TPR}}_{\mathrm{IPW}}(c) = \frac{\hat{\beta}_{1,\mathrm{IPW}}}{\hat{\theta}_{\mathrm{IPW}}} = \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c) V_i D_i \hat{\pi}_i^{-1}}{\sum_{i=1}^{n} V_i D_i \hat{\pi}_i^{-1}}, \tag{2.4}$$

$$\widehat{\mathrm{FPR}}_{\mathrm{IPW}}(c) = \frac{\hat{\beta}_{0,\mathrm{IPW}}}{1-\hat{\theta}_{\mathrm{IPW}}} = \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c) V_i (1-D_i) \hat{\pi}_i^{-1}}{\sum_{i=1}^{n} V_i (1-D_i) \hat{\pi}_i^{-1}}.$$

Again, the estimates $\hat{\pi}_i$ need to be obtained by using parametric regression models such as logistic or probit models. Finally, SPE estimators are defined as:

$$\widehat{\mathrm{TPR}}_{\mathrm{SPE}}(c) = \frac{\hat{\beta}_{1,\mathrm{SPE}}}{\hat{\theta}_{\mathrm{SPE}}} = \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c)\left\{V_i D_i \hat{\pi}_i^{-1} - (V_i \hat{\pi}_i^{-1} - 1)\hat{\rho}_i\right\}}{\sum_{i=1}^{n}\left\{V_i D_i \hat{\pi}_i^{-1} - (V_i \hat{\pi}_i^{-1} - 1)\hat{\rho}_i\right\}}, \tag{2.5}$$

$$\widehat{\mathrm{FPR}}_{\mathrm{SPE}}(c) = \frac{\hat{\beta}_{0,\mathrm{SPE}}}{1-\hat{\theta}_{\mathrm{SPE}}} = \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c)\left\{V_i (1-D_i) \hat{\pi}_i^{-1} - (V_i \hat{\pi}_i^{-1} - 1)(1-\hat{\rho}_i)\right\}}{\sum_{i=1}^{n}\left\{V_i (1-D_i) \hat{\pi}_i^{-1} - (V_i \hat{\pi}_i^{-1} - 1)(1-\hat{\rho}_i)\right\}}.$$

Alonzo and Pepe [3] showed that SPE estimators are doubly robust because they are consistent if either the π_i's or the ρ_i's are consistently estimated, i.e. provided that either of the two models employed for the disease and verification processes is correctly specified. However, it is worth noting that SPE estimates may not be range-respecting, i.e., they could fall outside the interval $(0, 1)$. This happens because the quantities $\left\{V_i D_i \hat{\pi}_i^{-1} - (V_i \hat{\pi}_i^{-1} - 1)\hat{\rho}_i\right\}$ or $\left\{V_i (1-D_i) \hat{\pi}_i^{-1} - (V_i \hat{\pi}_i^{-1} - 1)(1-\hat{\rho}_i)\right\}$ can be negative.

For each of the above methods, an estimated bias-corrected ROC curve can be obtained by plotting $\widehat{\mathrm{TPR}}(c)$ versus $\widehat{\mathrm{FPR}}(c)$ for all cut points c.

3. Proposal

Consider now a three–class problem. We model the disease status by a trinomial random vector $\mathcal{D} = (D_1, D_2, D_3)$, such that D_k is a Bernoulli random variable

having mean $\theta_k = \Pr(D_k = 1)$, with $\theta_1+\theta_2+\theta_3 = 1$. Let $\beta_{jk} = \Pr(T \geq c_j, D_k = 1)$ with $j = 1, 2$ and $k = 1, 2, 3$. In this notation,

$$\begin{aligned}
\mathrm{TCF}_1(c_1) &= 1 - \frac{\Pr(T \geq c_1, D_1 = 1)}{\Pr(D_1 = 1)} = 1 - \frac{\beta_{11}}{\theta_1}, \\
\mathrm{TCF}_2(c_1, c_2) &= \frac{\Pr(T \geq c_1, D_2 = 1) - \Pr(T \geq c_2, D_2 = 1)}{\Pr(D_2 = 1)} = \frac{\beta_{12} - \beta_{22}}{\theta_2}, \\
\mathrm{TCF}_3(c_2) &= \frac{\Pr(T \geq c_2, D_3 = 1)}{\Pr(D_3 = 1)} = \frac{\beta_{23}}{\theta_3}. \qquad (3.1)
\end{aligned}$$

When all subjects are verified, the nonparametric estimators of $\mathrm{TCF}_1, \mathrm{TCF}_2$ and TCF_3 are given by

$$\begin{aligned}
\widehat{\mathrm{TCF}}_1(c_1) &= 1 - \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c_1) D_{1i}}{\sum_{i=1}^{n} D_{1i}}, \\
\widehat{\mathrm{TCF}}_2(c_1, c_2) &= \frac{\sum_{i=1}^{n} \{\mathrm{I}(T_i \geq c_1) - \mathrm{I}(T_i \geq c_2)\} D_{2i}}{\sum_{i=1}^{n} D_{2i}}, \\
\widehat{\mathrm{TCF}}_3(c_2) &= \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c_2) D_{3i}}{\sum_{i=1}^{n} D_{3i}}.
\end{aligned}$$

In the presence of verification bias, we propose four estimators for $\mathrm{TCF}_1(c_1)$, $\mathrm{TCF}_2(c_1, c_2)$ and $\mathrm{TCF}_3(c_2)$. The proposed estimators work under the MAR assumption and are based on the estimation of the quantities $\theta_1, \theta_2, \beta_{11}, \beta_{12}, \beta_{22}$ and β_{23}. They can be seen as an extension of estimators reviewed in Subsection 2.2. In expressions (2.1) and (3.1), we note that parameters θ and θ_k, so as β_1 and β_{jk}, play, in essence, a similar role. Therefore, estimates of θ_k and β_{jk} can be obtained by mimicking what was done in the two-class problem.

3.1. Full imputation

For each $j = 1, 2$ and $k = 1, 2, 3$, the FI estimators of θ_k and β_{jk} are obtained as

$$\hat{\theta}_{k,\mathrm{FI}} = \widehat{\Pr}(D_k = 1) = \frac{1}{n} \sum_{i=1}^{n} \hat{\rho}_{ki}, \qquad (3.2)$$

$$\hat{\beta}_{jk,\mathrm{FI}} = \widehat{\Pr}(T \geq c_j, D_k = 1) = \frac{1}{n} \sum_{i=1}^{n} \mathrm{I}(T_i \geq c_j) \hat{\rho}_{ki}, \qquad (3.3)$$

where $\hat{\rho}_{ki}$ is an estimate of $\rho_{ki} = \Pr(D_{ki} = 1 | T_i, A_i)$ given by some suitable model, such as the multinomial logistic or probit regression model, applied to the verified sample units. Note that in what follows we will perform estimation within the framework of maximum likelihood. Therefore, the FI estimator

$\widehat{\mathrm{TCF}}_{1,\mathrm{FI}}(c_1)$, $\widehat{\mathrm{TCF}}_{2,\mathrm{FI}}(c_1,c_2)$ and $\widehat{\mathrm{TCF}}_{3,\mathrm{FI}}(c_2)$ are

$$\begin{aligned}
\widehat{\mathrm{TCF}}_{1,\mathrm{FI}}(c_1) &= 1-\frac{\sum_{i=1}^{n}\mathrm{I}(T_i\geq c_1)\hat{\rho}_{1i}}{\sum_{i=1}^{n}\hat{\rho}_{1i}},\\
\widehat{\mathrm{TCF}}_{2,\mathrm{FI}}(c_1,c_2) &= \frac{\sum_{i=1}^{n}\mathrm{I}(c_1\leq T_i<c_2)\hat{\rho}_{2i}}{\sum_{i=1}^{n}\hat{\rho}_{2i}},\\
\widehat{\mathrm{TCF}}_{3,\mathrm{FI}}(c_2) &= \frac{\sum_{i=1}^{n}\mathrm{I}(T_i\geq c_2)\hat{\rho}_{3i}}{\sum_{i=1}^{n}\hat{\rho}_{3i}}.
\end{aligned}$$

It is worth noting that estimates $\hat{\theta}_{k,\mathrm{FI}}$ and $\hat{\beta}_{jk,\mathrm{FI}}$ in (3.2) and (3.3) are the solutions of the estimating equations

$$\sum_{i=1}^{n}(\hat{\rho}_{ki}-\theta_k)=0 \quad \text{and} \quad \sum_{i=1}^{n}\{\mathrm{I}(T_i\geq c_j)\hat{\rho}_{ki}-\beta_{jk}\}=0. \tag{3.4}$$

3.2. Mean score imputation

By inspection of (2.3), we get the MSI estimators of θ_k, $k=1,2,3$, as follows

$$\hat{\theta}_{k,\mathrm{MSI}}=\widehat{\Pr}(D_k=1)=\frac{1}{n}\sum_{i=1}^{n}\left[V_iD_{ki}+(1-V_i)\hat{\rho}_{ki}\right].$$

The estimators of β_{jk} are given by

$$\hat{\beta}_{jk,\mathrm{MSI}}=\widehat{\Pr}(T\geq c_j,D_k=1)=\frac{1}{n}\sum_{i=1}^{n}\mathrm{I}(T_i\geq c_j)\left[V_iD_{ki}+(1-V_i)\hat{\rho}_{ki}\right].$$

Then, the MSI estimators of $\mathrm{TCF}_1(c_1)$, $\mathrm{TCF}_2(c_1,c_2)$ and $\mathrm{TCF}_3(c_2)$ are:

$$\begin{aligned}
\widehat{\mathrm{TCF}}_{1,\mathrm{MSI}}(c_1) &= 1-\frac{\sum_{i=1}^{n}\mathrm{I}(T_i\geq c_1)\left[V_iD_{1i}+(1-V_i)\hat{\rho}_{1i}\right]}{\sum_{i=1}^{n}\left[V_iD_{1i}+(1-V_i)\hat{\rho}_{1i}\right]},\\
\widehat{\mathrm{TCF}}_{2,\mathrm{MSI}}(c_1,c_2) &= \frac{\sum_{i=1}^{n}\mathrm{I}(c_1\leq T_i<c_2)\left[V_iD_{2i}+(1-V_i)\hat{\rho}_{2i}\right]}{\sum_{i=1}^{n}\left[V_iD_{2i}+(1-V_i)\hat{\rho}_{2i}\right]},\\
\widehat{\mathrm{TCF}}_{3,\mathrm{MSI}}(c_2) &= \frac{\sum_{i=1}^{n}\mathrm{I}(T_i\geq c_2)\left[V_iD_{3i}+(1-V_i)\hat{\rho}_{3i}\right]}{\sum_{i=1}^{n}\left[V_iD_{3i}+(1-V_i)\hat{\rho}_{3i}\right]}.
\end{aligned}$$

Again, we can obtain $\hat{\theta}_k$ and $\hat{\beta}_{jk}$ as solution of the estimating equations

$$\sum_{i=1}^{n} \{V_i(D_{ki} - \theta_k) + (1 - V_i)(\hat{\rho}_{ki} - \theta_k)\} = 0, \tag{3.5}$$

$$\sum_{i=1}^{n} \{V_i(\mathrm{I}(T_i \geq c_j)D_{ki} - \beta_{jk}) + (1 - V_i)(\mathrm{I}(T_i \geq c_j)\hat{\rho}_{ki} - \beta_{jk})\} = 0. \tag{3.6}$$

3.3. Inverse probability weighted

From the IPW estimators of β_1 and θ in (2.4), we derive, by analogy,

$$\begin{aligned}
\hat{\theta}_{k,\mathrm{IPW}} &= \widehat{\Pr}(D_k = 1) = \frac{\sum_{i=1}^{n} V_i \hat{\pi}_i^{-1} D_{ki}}{\sum_{i=1}^{n} V_i \hat{\pi}_i^{-1}}, \\
\hat{\beta}_{jk,\mathrm{IPW}} &= \widehat{\Pr}(T \geq c_j, D_k = 1) = \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c_j) V_i \hat{\pi}_i^{-1} D_{ki}}{\sum_{i=1}^{n} V_i \hat{\pi}_i^{-1}}.
\end{aligned}$$

The estimates $\hat{\pi}_i$ are obtained in the same way as in the two-class case, i.e., by using parametric regression models such as logistic or probit models. Again, in what follows we will employ maximum likelihood estimation. Then, the IPW estimates $\widehat{\mathrm{TCF}}_{1,\mathrm{IPW}}(c_1), \widehat{\mathrm{TCF}}_{2,\mathrm{IPW}}(c_1, c_2)$ and $\widehat{\mathrm{TCF}}_{3,\mathrm{IPW}}(c_2)$ are

$$\begin{aligned}
\widehat{\mathrm{TCF}}_{1,\mathrm{IPW}}(c_1) &= 1 - \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c_1) V_i \hat{\pi}_i^{-1} D_{1i}}{\sum_{i=1}^{n} V_i \hat{\pi}_i^{-1} D_{1i}}, \\
\widehat{\mathrm{TCF}}_{2,\mathrm{IPW}}(c_1, c_2) &= \frac{\sum_{i=1}^{n} \mathrm{I}(c_1 \leq T_i < c_2) V_i \hat{\pi}_i^{-1} D_{2i}}{\sum_{i=1}^{n} V_i \hat{\pi}_i^{-1} D_{2i}}, \\
\widehat{\mathrm{TCF}}_{3,\mathrm{IPW}}(c_2) &= \frac{\sum_{i=1}^{n} \mathrm{I}(T_i \geq c_2) V_i \hat{\pi}_i^{-1} D_{3i}}{\sum_{i=1}^{n} V_i \hat{\pi}_i^{-1} D_{3i}},
\end{aligned}$$

and the estimating equations corresponding to $\hat{\theta}_{k,\mathrm{IPW}}$ and $\hat{\beta}_{jk,\mathrm{IPW}}$ are

$$\begin{aligned}
\sum_{i=1}^{n} V_i \hat{\pi}_i^{-1} (D_{ki} - \theta_k) &= 0, \qquad (3.7) \\
\sum_{i=1}^{n} V_i \hat{\pi}_i^{-1} (\mathrm{I}(T_i \geq c_j) D_{ki} - \beta_{jk}) &= 0. \qquad (3.8)
\end{aligned}$$

Note that the IPW estimators only use verified subjects.

3.4. Semiparametric efficient

Similarly to three previous cases, the SPE estimators of β_{jk} and θ_k are derived in analogy to $\hat{\beta}_{1,\mathrm{SPE}}$ and $\hat{\theta}_{\mathrm{SPE}}$ in (2.5), i.e,

$$\hat{\theta}_{k,\mathrm{SPE}} = \frac{1}{n}\sum_{i=1}^{n}\left\{V_i D_{ki}\hat{\pi}_i^{-1} - \hat{\rho}_{ki}(V_i\hat{\pi}_i^{-1}-1)\right\}, \tag{3.9}$$

$$\hat{\beta}_{jk,\mathrm{SPE}} = \frac{1}{n}\sum_{i=1}^{n}\mathrm{I}(T_i \geq c_j)\left\{V_i D_{ki}\hat{\pi}_i^{-1} - \hat{\rho}_{ki}(V_i\hat{\pi}_i^{-1}-1)\right\}. \tag{3.10}$$

Therefore, we obtain

$$\widehat{\mathrm{TCF}}_{1,\mathrm{SPE}}(c_1) = 1 - \frac{\sum_{i=1}^{n}\mathrm{I}(T_i\geq c_1)\left\{V_i D_{1i}\hat{\pi}_i^{-1} - \hat{\rho}_{1i}(V_i\hat{\pi}_i^{-1}-1)\right\}}{\sum_{i=1}^{n}\left\{V_i D_{1i}\hat{\pi}_i^{-1} - \hat{\rho}_{1i}(V_i\hat{\pi}_i^{-1}-1)\right\}},$$

$$\widehat{\mathrm{TCF}}_{2,\mathrm{SPE}}(c_1,c_2) = \frac{\sum_{i=1}^{n}\mathrm{I}(c_1\leq T_i< c_2)\left\{V_i D_{2i}\hat{\pi}_i^{-1} - \hat{\rho}_{2i}(V_i\hat{\pi}_i^{-1}-1)\right\}}{\sum_{i=1}^{n}\left\{V_i D_{2i}\hat{\pi}_i^{-1} - \hat{\rho}_{2i}(V_i\hat{\pi}_i^{-1}-1)\right\}},$$

$$\widehat{\mathrm{TCF}}_{3,\mathrm{SPE}}(c_2) = \frac{\sum_{i=1}^{n}\mathrm{I}(T_i\geq c_2)\left\{V_i D_{3i}\hat{\pi}_i^{-1} - \hat{\rho}_{3i}(V_i\hat{\pi}_i^{-1}-1)\right\}}{\sum_{i=1}^{n}\left\{V_i D_{3i}\hat{\pi}_i^{-1} - \hat{\rho}_{3i}(V_i\hat{\pi}_i^{-1}-1)\right\}}.$$

The estimates $\hat{\theta}_{k,\mathrm{SPE}}$ and $\hat{\beta}_{jk,\mathrm{SPE}}$ solve the estimating equations

$$\sum_{i=1}^{n}\left\{\frac{V_i}{\hat{\pi}_i}\left[\mathrm{I}(T_i\geq c_j)D_{ki} - \beta_{jk}\right] - \frac{V_i-\hat{\pi}_i}{\hat{\pi}_i}\left[\mathrm{I}(T_i\geq c_j)\hat{\rho}_{ki} - \beta_{jk}\right]\right\} = 0, \tag{3.11}$$

$$\sum_{i=1}^{n}\left\{\frac{V_i}{\hat{\pi}_i}\left(D_{ki}-\theta_k\right) - \frac{V_i-\hat{\pi}_i}{\hat{\pi}_i}\left(\hat{\rho}_{ki}-\theta_k\right)\right\} = 0. \tag{3.12}$$

3.5. Asymptotic distribution theory

The parameters of interest $\mathrm{TCF}_1(c_1), \mathrm{TCF}_2(c_1,c_2)$ and $\mathrm{TCF}_3(c_2)$ are functions of θ_1, θ_2, β_{11}, β_{12}, β_{22}, β_{23} and $\tau = (\tau_\rho^\top, \tau_\pi^\top)^\top$, where τ is the vector of parameters of the models used to estimate $\rho = (\rho_1,\rho_2)^\top$, or π, or both. Let us denote $\alpha = (\theta_1,\theta_2,\beta_{11},\beta_{12},\beta_{22},\beta_{23},\tau^\top)^\top$. The estimators (FI, MSI, IPW, SPE) of α are obtained by solving suitable estimating equations. Hence, we use results in Alonzo, Pepe and Lumley [2] and Alonzo and Pepe [3] to give consistency and asymptotic normality of the proposed bias–corrected estimators.

According to equations (3.4), (3.5), (3.6), (3.7), (3.8), (3.11) and (3.12), let $G_*^{\theta_s}(\alpha) = \sum_{i=1}^{n} g_{i,*}^{\theta_s}(\alpha)$ and $G_*^{\beta_{jk}}(\alpha) = \sum_{i=1}^{n} g_{i,*}^{\beta_{jk}}(\alpha)$ be the estimating functions for θ_s and β_{jk}, with $k = 1,2,3$, s and $j = 1,2$, for one of the four previously introduced approaches (the star indicates FI, MSI, IPW, SPE). We assume that

$\hat{\tau}$ is the solution to a classic set of estimating equations of the form $G^{\tau}(\alpha) = \sum_{i=1}^{n} g_i^{\tau}(\alpha) = 0$. Specifically, we will employ classic score equations derived from: a multinomial logistic regression model for estimation of the disease process; and from a logistic or probit regression model for estimation of the verification process. The estimate $\hat{\alpha}_*$ of α is then obtained by solving $G_*(\alpha) = \sum_{i=1}^{n} g_{i,*}(\alpha) = 0$, where $g_{i,*}(\alpha) = \left(g_{i,*}^{\theta_1}(\alpha), g_{i,*}^{\theta_2}(\alpha), g_{i,*}^{\beta_{11}}(\alpha), g_{i,*}^{\beta_{12}}(\alpha), g_{i,*}^{\beta_{22}}(\alpha), g_{i,*}^{\beta_{23}}(\alpha), g_i^{\tau}(\alpha)^{\top}\right)^{\top}$.

Let $\alpha_0 = (\theta_{10}, \theta_{20}, \beta_{110}, \beta_{120}, \beta_{220}, \beta_{230}, \tau_0^{\top})^{\top}$ be the true value of α. We assume that

(A1) $\mathcal{D}$ is missing at random (MAR);
(A2) the data $(\mathcal{D}_i, T_i, A_i, V_i)$ are iid;
(A3) $(T, A)^{\top}$ is a bounded random vector;
(A4) $\mathbb{E}\left[\frac{\partial}{\partial \alpha^{\top}} g_{i,*}(\alpha_0)\right]$ is negative definite;
(A5) ρ_{ki} and π_i are bounded away from 0.

We consider also the following standard regularity conditions.

(C1) $g_{i,*}(\alpha_0)$ are iid and $\mathbb{E}\{g_{i,*}(\alpha_0)\} = 0$.
(C2) Elements of $G_*(\alpha)$, $\frac{\partial}{\partial \alpha^{\top}} G_*(\alpha)$, and $\frac{\partial^2}{\partial \alpha \partial \alpha^{\top}} G_*(\alpha)$ exist in a bounded δ-neighborhood of α_0, $N_{\delta}(\alpha_0)$.
(C3) $g_{i,*}(\alpha)$, $\frac{\partial}{\partial \alpha^{\top}} g_{i,*}(\alpha)$, and $\frac{\partial^2}{\partial \alpha \partial \alpha^{\top}} g_{i,*}(\alpha)$ are uniformly bounded in $N_{\delta}(\alpha_0)$.

Under the assumptions (A1)–(A5) and conditions (C1)–(C3), we obtain the asymptotic results summarized in the following theorem.

Theorem 3.1. *Let* $\mathrm{TCF}_{10}(c_1), \mathrm{TCF}_{20}(c_1, c_2), \mathrm{TCF}_{30}(c_2)$ *be the true parameter values. The FI, MSI, IPW or SPE bias-corrected estimators* $\widehat{\mathrm{TCF}}_{1,*}(c_1)$, $\widehat{\mathrm{TCF}}_{2,*}(c_1, c_2)$ *and* $\widehat{\mathrm{TCF}}_{3,*}(c_2)$ *are consistent. Furthermore,*

$$\sqrt{n}\left[\begin{pmatrix} \widehat{\mathrm{TCF}}_{1,*}(c_1) \\ \widehat{\mathrm{TCF}}_{2,*}(c_1, c_2) \\ \widehat{\mathrm{TCF}}_{3,*}(c_2) \end{pmatrix} - \begin{pmatrix} \mathrm{TCF}_{10}(c_1) \\ \mathrm{TCF}_{20}(c_1, c_2) \\ \mathrm{TCF}_{30}(c_2) \end{pmatrix}\right] \xrightarrow{d} \mathcal{N}_3\left(\mathbf{0}, \Xi\right), \tag{3.13}$$

where

$$\Xi = \frac{\partial h(\alpha_0)}{\partial \alpha^{\top}} \Sigma \frac{\partial h^{\top}(\alpha_0)}{\partial \alpha^{\top}},$$

with $h(\alpha) = \left(1 - \frac{\beta_{11}}{\theta_1}, \frac{\beta_{12} - \beta_{22}}{\theta_2}, \frac{\beta_{23}}{1 - (\theta_1 + \theta_2)}\right)^{\top}$ *and*

$$\Sigma = \left[\mathbb{E}\left\{\frac{\partial}{\partial \alpha^{\top}} g_{i,*}(\alpha_0)\right\}\right]^{-1} \mathbb{C}\mathrm{ov}\{g_{i,*}(\alpha_0)\} \left[\mathbb{E}\left\{\frac{\partial}{\partial \alpha^{\top}} g_{i,*}^{\top}(\alpha_0)\right\}\right]^{-1}.$$

Proof. We apply Theorem 1 and Theorem 2 of [2]. Under assumptions (A1)–(A5) and conditions (C1)–(C3), $\hat{\alpha}_*$ is consistent and $\sqrt{n}\left(\hat{\alpha}_* - \alpha_0\right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Xi)$. Thus, the estimators $\widehat{\mathrm{TCF}}_{1,*}(c_1) = 1 - \hat{\beta}_{11}/\hat{\theta}_1, \widehat{\mathrm{TCF}}_{2,*}(c_1, c_2) = (\hat{\beta}_{12} - \hat{\beta}_{22})/\hat{\theta}_2$

and $\widehat{\mathrm{TCF}}_{3,*}(c_2) = \hat{\beta}_{23}/(1-(\hat{\theta}_1+\hat{\theta}_2))$ are consistent for the true $\mathrm{TCF}_{10}(c_1)$, $\mathrm{TCF}_{20}(c_1,c_2)$ and $\mathrm{TCF}_{30}(c_2)$ and, by and application of the multivariate delta method, result (3.13) follows. In Appendix A, we check conditions (C1)–(C3) for each estimator, i.e., FI, MSI, IPW and SPE, under assumptions (A1)–(A5). This is done when a multinomial logistic regression model is used for the estimation of the disease process and a logistic or probit regression model is used for the estimation of the verification process. □

The above theorem gives a general result for all estimates, i.e., FI, MSI, IPW and SPE. In Appendix, the explicit form of the asymptotic variance–covariance matrix is obtained. In practice, the variance–covariance matrix Σ is replaced by a consistent estimate $\hat{\Sigma}$

$$\hat{\Sigma} = n\left[\sum_{i=1}^{n}\frac{\partial}{\partial\alpha^{\top}}g_{i,*}(\hat{\alpha})\right]^{-1}\left[\sum_{i=1}^{n}g_{i,*}(\hat{\alpha})g_{i,*}(\hat{\alpha})^{\top}\right]\left[\sum_{i=1}^{n}\frac{\partial}{\partial\alpha^{\top}}g_{i,*}^{\top}(\hat{\alpha})\right]^{-1}.$$

It is worth noting that SPE estimators of θ_k and β_{jk} in (3.9) and (3.10), will inherit the double robustness property of $\hat{\theta}_{\mathrm{SPE}}$ and $\hat{\beta}_{1,\mathrm{SPE}}$ in (2.5). That is, $\hat{\theta}_{k,\mathrm{SPE}}$ and $\hat{\beta}_{jk,\mathrm{SPE}}$ remain consistent if only one of the disease model $P(D_k = 1|T,A)$ or the verification model $P(V=1|T,A)$ is correctly specified in the estimation process; they are inconsistent if both models are misspecified. Clearly, this property holds also for the estimators $\widehat{\mathrm{TCF}}_{1,\mathrm{SPE}}(c_1)$, $\widehat{\mathrm{TCF}}_{2,\mathrm{SPE}}(c_1,c_2)$ and $\widehat{\mathrm{TCF}}_{3,\mathrm{SPE}}(c_2)$.

3.6. VUS estimation

Let μ be the volume under the ROC surface (VUS) of T. A straightforward calculation (Nakas and Yiannoutsos [13]) shows that

$$\begin{aligned}\mu &= \Pr\left(T_i < T_\ell < T_r | D_{1i}=1, D_{2\ell}=1, D_{3r}=1\right)\\ &\quad + \frac{1}{2}\Pr\left(T_i < T_\ell = T_r | D_{1i}=1, D_{2\ell}=1, D_{3r}=1\right)\\ &\quad + \frac{1}{2}\Pr\left(T_i = T_\ell < T_r | D_{1i}=1, D_{2\ell}=1, D_{3r}=1\right)\\ &\quad + \frac{1}{6}\Pr\left(T_i = T_\ell = T_r | D_{1i}=1, D_{2\ell}=1, D_{3r}=1\right)\end{aligned}$$

or, equivalently,

$$\mu = \frac{\mathbb{E}\left(D_{1i}D_{2\ell}D_{3r}\mathrm{I}_{i\ell r}\right)}{\mathbb{E}\left(D_{1i}D_{2\ell}D_{3r}\right)},$$

where $\mathrm{I}_{i\ell r} = \mathrm{I}(T_i < T_\ell < T_r) + \frac{1}{2}\mathrm{I}(T_i < T_\ell = T_r) + \frac{1}{2}\mathrm{I}(T_i = T_\ell < T_r) + \frac{1}{6}\mathrm{I}(T_i = T_\ell = T_r)$. Then, in the absence of missing data, a natural nonparametric

estimator $\hat{\mu}$ of μ is given by

$$\hat{\mu} = \frac{\sum_{i=1}^{n} \sum_{\ell=1,\ell\neq i}^{n} \sum_{\substack{r=1 \\ r\neq\ell, r\neq i}}^{n} \mathrm{I}_{i\ell r} D_{1i} D_{2\ell} D_{3r}}{\sum_{i=1}^{n} \sum_{\ell=1,\ell\neq i}^{n} \sum_{\substack{r=1 \\ r\neq\ell, r\neq i}}^{n} D_{1i} D_{2\ell} D_{3r}}. \tag{3.14}$$

When the disease status is missing for some of the subjects, verification bias–corrected estimators of VUS, can be obtained by using suitable estimates of quantities D_{1i}, D_{2i} and D_{3i} in (3.14). More precisely, FI, MSI, IPW and SPE estimators of VUS take the form

$$\hat{\mu}_* = \frac{\sum_{i=1}^{n} \sum_{\ell=1,\ell\neq i}^{n} \sum_{\substack{r=1 \\ r\neq\ell, r\neq i}}^{n} \mathrm{I}_{i\ell r} \tilde{D}_{1i,*} \tilde{D}_{2\ell,*} \tilde{D}_{3r,*}}{\sum_{i=1}^{n} \sum_{\ell=1,\ell\neq i}^{n} \sum_{\substack{r=1 \\ r\neq\ell, r\neq i}}^{n} \tilde{D}_{1i,*} \tilde{D}_{2\ell,*} \tilde{D}_{3r,*}},$$

where the star again stands for FI, MSI, IPW, SPE, and

$$\tilde{D}_{ki,\mathrm{FI}} = \hat{\rho}_{ki}, \quad \tilde{D}_{ki,\mathrm{MSI}} = V_i D_{ki} + (1 - V_i)\hat{\rho}_{ki}, \quad \tilde{D}_{ki,\mathrm{IPW}} = V_i D_{ki} \hat{\pi}_i^{-1},$$
$$\tilde{D}_{ki,\mathrm{SPE}} = V_i D_{ki} \hat{\pi}_i^{-1} - \hat{\rho}_{ki}(V_i \hat{\pi}_i^{-1} - 1) \qquad (k = 1, 2, 3).$$

As for the estimators of the TCFs, under the MAR assumption and certain suitable regularity conditions, we can establish consistency and asymptotic normality of the above given bias-corrected VUS estimators (proof available from the authors). Moreover, the asymptotic variance of $\hat{\mu}_*$, i.e. the variance of $\sqrt{n}\,(\hat{\mu}_* - \mu_0)$, can be consistently estimated by

$$\frac{\frac{1}{n-1} \sum_{i=1}^{n} \hat{Q}_{i,*}^2(\hat{\mu}_*, \hat{\tau})}{\hat{\theta}_{1,*}^2 \hat{\theta}_{2,*}^2 \hat{\theta}_{3,*}^2},$$

where

$$\begin{aligned}
\hat{Q}_{i,*}(\hat{\mu}_*, \hat{\tau}) = &- \left\{ \frac{1}{(n-1)(n-2)} \sum_{i=1}^{n} \sum_{\substack{\ell=i \\ \ell\neq i}}^{n} \sum_{\substack{r=1 \\ r\neq\ell, r\neq i}}^{n} \left. \frac{\partial G_{i\ell r,*}^{\top}(\hat{\mu}_*, \tau_\rho, \hat{\tau}_\pi)}{\partial \tau_\rho} \right|_{\tau_\rho = \hat{\tau}_\rho} \right\} \\
&\times \left\{ \sum_{i=1}^{n} \left. \frac{\partial g_i^{\tau_\rho}}{\partial \tau_\rho^{\top}} \right|_{\tau_\rho = \hat{\tau}_\rho} \right\}^{-1} \left. g_i^{\tau_\rho} \right|_{\tau_\rho = \hat{\tau}_\rho} \\
&- \left\{ \frac{1}{(n-1)(n-2)} \sum_{i=1}^{n} \sum_{\substack{\ell=i \\ \ell\neq i}}^{n} \sum_{\substack{r=1 \\ r\neq\ell, r\neq i}}^{n} \left. \frac{\partial G_{i\ell r,*}^{\top}(\hat{\mu}_*, \hat{\tau}_\rho, \tau_\pi)}{\partial \tau_\pi} \right|_{\tau_\pi = \hat{\tau}_\pi} \right\} \\
&\times \left\{ \sum_{i=1}^{n} \left. \frac{\partial g_i^{\tau_\pi}}{\partial \tau_\pi^{\top}} \right|_{\tau_\pi = \hat{\tau}_\pi} \right\}^{-1} \left. g_i^{\tau_\pi} \right|_{\tau_\pi = \hat{\tau}_\pi}
\end{aligned}$$

$$+ \frac{1}{(n-1)(n-2)} \sum_{\substack{\ell=1 \\ \ell \neq i}}^{n} \sum_{\substack{r=1 \\ r \neq i, r \neq \ell}}^{n} \Big\{ G_{i\ell r,*}(\hat{\mu}_*, \hat{\tau}_\rho, \hat{\tau}_\pi) + G_{\ell i r,*}(\hat{\mu}_*, \hat{\tau}_\rho, \hat{\tau}_\pi) + G_{r\ell i,*}(\hat{\mu}_*, \hat{\tau}_\rho, \hat{\tau}_\pi) \Big\}, \tag{3.15}$$

with

$$\begin{aligned}
G_{i\ell r,\mathrm{FI}}(\mu, \tau_\rho, \tau_\pi) &= \rho_{1i}(\tau_\rho)\rho_{2\ell}(\tau_\rho)\rho_{3r}(\tau_\rho)\,(I_{i\ell r} - \mu)\,, \\
G_{i\ell r,\mathrm{MSI}}(\mu, \tau_\rho, \tau_\pi) &= D_{1i,\mathrm{MSI}}(\tau_\rho)D_{2\ell,\mathrm{MSI}}(\tau_\rho)D_{3r,\mathrm{MSI}}(\tau_\rho)\,(I_{i\ell r} - \mu)\,, \\
G_{i\ell r,\mathrm{IPW}}(\mu, \tau_\rho, \tau_\pi) &= \frac{V_i V_\ell V_r}{\pi_i(\tau_\pi)\pi_\ell(\tau_\pi)\pi_r(\tau_\pi)} D_{1i}D_{2\ell}D_{3r}\,(I_{i\ell r} - \mu)\,, \\
G_{i\ell r,\mathrm{SPE}}(\mu, \tau_\rho, \tau_\pi) &= D_{1i,\mathrm{SPE}}(\tau_\rho, \tau_\pi)D_{2\ell,\mathrm{SPE}}(\tau_\rho, \tau_\pi)D_{3r,\mathrm{SPE}}(\tau_\rho, \tau_\pi)\,(I_{i\ell r} - \mu)\,,
\end{aligned}$$

and

$$\begin{aligned}
D_{ki,\mathrm{MSI}}(\tau_\rho) &= V_i D_{ki} + (1 - V_i)\rho_{ki}(\tau_\rho), \\
D_{ki,\mathrm{SPE}}(\tau_\rho, \tau_\pi) &= V_i D_{ki}\pi_i^{-1}(\tau_\pi) - \rho_{ki}(\tau_\rho)(V_i\pi_i^{-1}(\tau_\pi) - 1),
\end{aligned}$$

for $k = 1, 2, 3$. In (3.15), the functions $g_i^{\tau_\rho}(\cdot)$ and $g_i^{\tau_\pi}(\cdot)$ are the elements of the functions $g_i^{\tau}(\cdot)$ in the estimating function $G^{\tau}(\cdot)$ for the parameters of the models adopted for the disease and the verification processes. See the Appendix A for their specification when the models chosen are, the multinomial logistic regression and the logistic (or probit) model, respectively.

4. Simulation studies

In this section, the ability of FI, MSI, IPW and SPE methods to estimate TCF_1, TCF_2 and TCF_3 are evaluated by using Monte Carlo experiments. Also, the square root of the estimates of the variances are compared with Monte Carlo and bootstrap standard deviations. Some simulation results concerning the behaviour of the estimators for the VUS are given in Appendix C.

Note that, the bias-corrected estimators of TCF_1, TCF_2 and TCF_3 require a parametric regression model to estimate $\rho_{ki} = \Pr(D_{ki} = 1|T_i, A_i)$, or $\pi_i = \Pr(V_i = 1|T_i, A_i)$, or both. A wrong specification of such models may affect the estimation. Therefore, in the simulation study we consider four scenarios:

(i) the disease model and the verification model are both correctly specified;
(ii) the verification model is misspecified;
(iii) the disease model is misspecified;
(iv) the disease model and the verification model are both misspecified.

All scenarios allow to evaluate the behavior of the proposed estimators in finite samples. In practice, we consider 5000 Monte Carlo replications, and three sample sizes, i.e., 250, 500 and 1000 in scenario (i) and a sample size equal to 1000 in scenarios (ii)–(iv). The choice of such sample size in scenarios (ii)–(iv) allows to dig up expected bad behaviors of the estimators under misspecification, when a great amount of information is available, i.e., in large samples.

4.1. Study 1

The true disease $\mathcal{D}$ is generated by a trinomial random vector (D_1, D_2, D_3), such that D_k is a Bernoulli random variable with mean θ_k, $k = 1, 2, 3$. We set $\theta_1 = 0.4, \theta_2 = 0.35$ and $\theta_3 = 0.25$. The continuous test results T and A are generated from the following conditional models

$$T, A|D_k \sim \mathcal{N}_2\left(\mu_k, \Lambda\right), \qquad k = 1, 2, 3,$$

where $\mu_k = (2k, k)^\top$. We consider three different values for Λ, specifically

$$\begin{pmatrix} 1.75 & 0.1 \\ 0.1 & 2.5 \end{pmatrix}, \qquad \begin{pmatrix} 2.5 & 1.5 \\ 1.5 & 2.5 \end{pmatrix}, \qquad \begin{pmatrix} 5.5 & 3 \\ 3 & 2.5 \end{pmatrix},$$

giving rise to a correlation between T and A equal to $0.36, 0.69$ and 0.84, respectively.

In this scenario -and also in the next one- we consider six pairs for cut points (c_1, c_2), i.e., $(2, 4), (2, 5), (2, 7), (4, 5), (4, 7)$ and $(5, 7)$. Since the conditional distribution of T given D_k is the normal distribution, the true values of TCF's are obtained as

$$\begin{aligned} \mathrm{TCF}_1(c_1) &= \Phi\left(\frac{c_1 - 2}{\sigma_{T|D}}\right), \\ \mathrm{TCF}_2(c_1, c_2) &= \Phi\left(\frac{c_2 - 4}{\sigma_{T|D}}\right) - \Phi\left(\frac{c_1 - 4}{\sigma_{T|D}}\right), \\ \mathrm{TCF}_3(c_2) &= 1 - \Phi\left(\frac{c_2 - 6}{\sigma_{T|D}}\right), \end{aligned}$$

where $\sigma_{T|D}$ denotes the entry in the 1-st row and 1-st column of Λ and $\phi(\cdot)$ and $\Phi(\cdot)$ are the density function and the cumulative distribution function of the standard normal random variable, respectively.

Under our data–generating process, the true conditional disease model is a multinomial logistic model

$$\Pr(D_k = 1|T, A) = \frac{\exp\left(\tau_{\rho_{1k}} + \tau_{\rho_{2k}} T + \tau_{\rho_{3k}} A\right)}{1 + \exp\left(\tau_{\rho_{11}} + \tau_{\rho_{21}} T + \tau_{\rho_{31}} A\right) + \exp\left(\tau_{\rho_{12}} + \tau_{\rho_{22}} T + \tau_{\rho_{32}} A\right)},$$

for suitable $\tau_{\rho_{1k}}, \tau_{\rho_{2k}}, \tau_{\rho_{3k}}$, where $k = 1, 2$. The verification status V is generated by the following model

$$\mathrm{logit}\left\{\Pr(V = 1|T, A)\right\} = 0.5 - 0.3T + 0.75A.$$

This choice corresponds to a verification rate of about 0.65. In this study, the FI, MSI, IPW and SPE estimators are computed under correct working models for both the disease and the verification status. Therefore, in particular, the conditional verification probabilities π_i are estimated from a logistic model for V given T and A.

Tables 1–3 and Tables 9–14 in Appendix B show Monte Carlo means, Monte Carlo standard deviations (MC.sd), the square roots of the variance estimated via asymptotic results (asy.sd) and bootstrap standard deviations (boot.sd) of $\widehat{\text{TCF}}_1$, $\widehat{\text{TCF}}_2$ and $\widehat{\text{TCF}}_3$. Here and in the following bootstrap estimates are obtained from 250 bootstrap replications. Overall, the estimators FI, MSI, IPW and SPE behave similarly in this scenario, with the IPW estimator showing a slightly bigger standard deviation. Simulation results, in this and in the following scenarios, also show that, excluding the SPE approach, bootstrap estimates of standard deviations are generally more accurate than estimates obtained via asymptotic theory.

4.2. Study 2

In this study, the true disease status $\mathcal{D}$ and the test results T and A are generated in the same way as in the first scenario. The true conditional verification process π, instead, is chosen to be the following function of T and A

$$\pi(T, A) = 0.35 + 0.3\text{I}\left(T > t^{(0.8)}\right) + 0.35\text{I}\left(A > a^{(0.8)}\right),$$

where $t^{(0.8)}$ and $a^{(0.8)}$ correspond to the 80-th percentile of distribution of T and A, respectively. In this case, the verification probabilities are 1 for subjects with $T > t^{(0.8)}$ and $A > a^{(0.8)}$; 0.7 for subjects with $T \leq t^{(0.8)}$ and $A > a^{(0.8)}$; 0.65 for subjects with $T > t^{(0.8)}$ and $A \leq a^{(0.8)}$; 0.35 otherwise. In our setting, the verification rate is approximately 0.48.

The aim in this scenario is to evaluate the behavior of the estimators, in particular that of IPW and SPE, under misspecification of the verification process. Therefore, $\hat{\pi}_i$ is estimated from a logistic regression model with V as the response and T as predictor, while $\hat{\rho}_{ki}$ is still obtained from the multinomial logistic model (similarly to the first scenario). Clearly, the model used for verification status is misspecified.

Table 4 and Tables 15–16 in Appendix B, show Monte Carlo means and standard deviations for the estimators of the true class fractions TCF_1, TCF_2 and TCF_3. Moreover, estimated standard deviations (via asymptotic theory) and bootstrap standard deviations are also presented. The results clearly show the effect of misspecification on IPW estimates, despite the high sample size. In particular, in terms of bias, the IPW method performs almost alway poorly, with high distortion in some cases (values highlighted in bold). On the other hand, the SPE estimator behaves well, due to its doubly robustness property.

4.3. Study 3

Starting from two independent random variables $Z_1 \sim \mathcal{N}(0, 0.5)$ and $Z_2 \sim \mathcal{N}(0, 0.5)$, the true conditional disease is generated by a trinomial random vector (D_1, D_2, D_3) such that

$$D_1 = \begin{cases} 1 & \text{if } Z_1 + Z_2 \leq h_1 \\ 0 & \text{otherwise} \end{cases}, \; D_2 = \begin{cases} 1 & \text{if } h_1 < Z_1 + Z_2 \leq h_2 \\ 0 & \text{otherwise} \end{cases},$$

TABLE 1. *Simulation results from 5000 replications when both models for ρ_k and π are correctly specified (Study 1) and the first value of Λ is considered. "True" denotes the true parameter value. Sample size = 250.*

	TCF_1	TCF_2	TCF_3	$MC.sd_1$	$MC.sd_2$	$MC.sd_3$	$asy.sd_1$	$asy.sd_2$	$asy.sd_3$	$boot.sd_1$	$boot.sd_2$	$boot.sd_3$
						cut-point = (2,4)						
True	0.5000	0.4347	0.9347									
FI	0.5008	0.4357	0.9342	0.0529	0.0472	0.0272	0.0486	0.0440	0.0512	0.0530	0.0488	0.0276
MSI	0.5007	0.4353	0.9341	0.0544	0.0536	0.0318	0.0500	0.0501	0.0538	0.0542	0.0542	0.0324
IPW	0.5017	0.4352	0.9341	0.0714	0.0721	0.0371	0.0687	0.0697	0.0398	0.0704	0.0711	0.0373
SPE	0.5008	0.4352	0.9343	0.0574	0.0648	0.0364	0.0562	0.0632	0.0340	0.0596	0.1497	0.0425
						cut-point = (2,5)						
True	0.5000	0.7099	0.7752									
FI	0.5008	0.7122	0.7756	0.0529	0.0464	0.0537	0.0486	0.0454	0.0618	0.0530	0.0467	0.0533
MSI	0.5007	0.7112	0.7747	0.0544	0.0511	0.0568	0.0500	0.0503	0.0644	0.0542	0.0514	0.0563
IPW	0.5017	0.7123	0.7739	0.0714	0.0683	0.0666	0.0687	0.0658	0.0704	0.0704	0.0677	0.0655
SPE	0.5008	0.7116	0.7751	0.0574	0.0619	0.0630	0.0562	0.0597	0.0603	0.0596	0.1219	0.1033
						cut-point = (2,7)						
True	0.5000	0.9230	0.2248									
FI	0.5008	0.9231	0.2229	0.0529	0.0236	0.0520	0.0486	0.0327	0.0437	0.0530	0.0243	0.0525
MSI	0.5007	0.9230	0.2230	0.0544	0.0285	0.0530	0.0500	0.0361	0.0447	0.0542	0.0287	0.0534
IPW	0.5017	0.9234	0.2216	0.0714	0.0376	0.0748	0.0687	0.0341	0.0706	0.0704	0.0368	0.0727
SPE	0.5008	0.9234	0.2236	0.0574	0.0361	0.0571	0.0562	0.0334	0.0559	0.0596	0.0474	0.4185
						cut-point = (4,5)						
True	0.9347	0.2752	0.7752									
FI	0.9350	0.2765	0.7756	0.0244	0.0408	0.0537	0.0224	0.0350	0.0618	0.0247	0.0415	0.0533
MSI	0.9351	0.2759	0.7747	0.0270	0.0467	0.0568	0.0247	0.0411	0.0644	0.0271	0.0467	0.0563
IPW	0.9356	0.2770	0.7739	0.0413	0.0690	0.0666	0.0343	0.0645	0.0704	0.0395	0.0663	0.0655
SPE	0.9356	0.2764	0.7751	0.0378	0.0587	0.0630	0.0333	0.0560	0.0603	0.0471	0.1566	0.1033
						cut-point = (4,7)						
True	0.9347	0.4883	0.2248									
FI	0.9350	0.4874	0.2229	0.0244	0.0523	0.0520	0.0224	0.0494	0.0437	0.0247	0.0537	0.0525
MSI	0.9351	0.4877	0.2230	0.0270	0.0559	0.0530	0.0247	0.0528	0.0447	0.0271	0.0567	0.0534
IPW	0.9356	0.4881	0.2216	0.0413	0.0741	0.0748	0.0343	0.0708	0.0706	0.0395	0.0723	0.0727
SPE	0.9356	0.4882	0.2236	0.0378	0.0661	0.0571	0.0333	0.0640	0.0559	0.0471	0.1745	0.4185
						cut-point = (5,7)						
True	0.9883	0.2132	0.2248									
FI	0.9880	0.2109	0.2229	0.0075	0.0432	0.0520	0.0066	0.0387	0.0437	0.0081	0.0436	0.0525
MSI	0.9881	0.2118	0.2230	0.0098	0.0460	0.0530	0.0075	0.0423	0.0447	0.0101	0.0468	0.0534
IPW	0.9885	0.2111	0.2216	0.0203	0.0634	0.0748	0.0097	0.0601	0.0706	0.0185	0.0625	0.0727
SPE	0.9883	0.2117	0.2236	0.0191	0.0569	0.0571	0.0117	0.0542	0.0559	0.0180	0.1382	0.4185

TABLE 2. *Simulation results from 5000 replications when both models for ρ_k and π are correctly specified (Study 1) and the second value of Λ is considered. "True" denotes the true parameter value. Sample size = 250.*

	TCF_1	TCF_2	TCF_3	$MC.sd_1$	$MC.sd_2$	$MC.sd_3$	$asy.sd_1$	$asy.sd_2$	$asy.sd_3$	$boot.sd_1$	$boot.sd_2$	$boot.sd_3$
						cut-point = (2,4)						
True	0.5000	0.3970	0.8970									
FI	0.4995	0.3966	0.8972	0.0502	0.0419	0.0357	0.0461	0.0375	0.0502	0.0506	0.0429	0.0362
MSI	0.4996	0.3966	0.8970	0.0519	0.0498	0.0409	0.0479	0.0463	0.0536	0.0522	0.0506	0.0410
IPW	0.5001	0.3972	0.8979	0.0659	0.0700	0.0523	0.0646	0.0677	0.0504	0.0658	0.0687	0.0510
SPE	0.4996	0.3968	0.8980	0.0565	0.0623	0.0508	0.0559	0.0617	0.0469	0.0576	0.1619	0.0502
						cut-point = (2,5)						
True	0.5000	0.6335	0.7365									
FI	0.4995	0.6340	0.7378	0.0502	0.0431	0.0580	0.0461	0.0410	0.0619	0.0506	0.0440	0.0580
MSI	0.4996	0.6335	0.7370	0.0519	0.0502	0.0617	0.0479	0.0485	0.0653	0.0522	0.0510	0.0616
IPW	0.5001	0.6330	0.7379	0.0659	0.0679	0.0733	0.0646	0.0660	0.0737	0.0658	0.0671	0.0721
SPE	0.4996	0.6335	0.7377	0.0565	0.0616	0.0686	0.0559	0.0610	0.0665	0.0576	0.1438	0.0686
						cut-point = (2,7)						
True	0.5000	0.8682	0.2635									
FI	0.4995	0.8679	0.2640	0.0502	0.0307	0.0559	0.0461	0.0333	0.0499	0.0506	0.0314	0.0558
MSI	0.4996	0.8680	0.2644	0.0519	0.0362	0.0588	0.0479	0.0387	0.0523	0.0522	0.0372	0.0580
IPW	0.5001	0.8678	0.2659	0.0659	0.0492	0.0695	0.0646	0.0472	0.0682	0.0658	0.0492	0.0690
SPE	0.4996	0.8684	0.2649	0.0565	0.0467	0.0615	0.0559	0.0451	0.0591	0.0576	0.0593	0.0610
						cut-point = (4,5)						
True	0.8970	0.2365	0.7365									
FI	0.8974	0.2374	0.7378	0.0284	0.0368	0.0580	0.0274	0.0318	0.0619	0.0288	0.0371	0.0580
MSI	0.8972	0.2369	0.7370	0.0320	0.0441	0.0617	0.0306	0.0395	0.0653	0.0320	0.0439	0.0616
IPW	0.8978	0.2358	0.7379	0.0377	0.0603	0.0733	0.0361	0.0574	0.0737	0.0372	0.0586	0.0721
SPE	0.8975	0.2368	0.7377	0.0364	0.0538	0.0686	0.0352	0.0519	0.0665	0.0363	0.2833	0.0686
						cut-point = (4,7)						
True	0.8970	0.4711	0.2635									
FI	0.8974	0.4713	0.2640	0.0284	0.0504	0.0559	0.0274	0.0467	0.0499	0.0288	0.0510	0.0558
MSI	0.8972	0.4714	0.2644	0.0320	0.0554	0.0588	0.0306	0.0525	0.0523	0.0320	0.0562	0.0580
IPW	0.8978	0.4706	0.2659	0.0377	0.0693	0.0695	0.0361	0.0677	0.0682	0.0372	0.0687	0.0690
SPE	0.8975	0.4716	0.2649	0.0364	0.0635	0.0615	0.0352	0.0627	0.0591	0.0363	0.1949	0.0610
						cut-point = (5,7)						
True	0.9711	0.2347	0.2635									
FI	0.9710	0.2339	0.2640	0.0121	0.0404	0.0559	0.0118	0.0369	0.0499	0.0127	0.0409	0.0558
MSI	0.9708	0.2345	0.2644	0.0165	0.0458	0.0588	0.0151	0.0431	0.0523	0.0167	0.0465	0.0580
IPW	0.9710	0.2348	0.2659	0.0203	0.0569	0.0695	0.0178	0.0556	0.0682	0.0201	0.0569	0.0690
SPE	0.9710	0.2348	0.2649	0.0201	0.0526	0.0615	0.0179	0.0521	0.0591	0.0199	0.1084	0.0610

TABLE 3. *Simulation results from 5000 replications when both models for ρ_k and π are correctly specified (Study 1) and the third value of Λ is considered. "True" denotes the true parameter value. Sample size = 250.*

	TCF_1	TCF_2	TCF_3	$MC.sd_1$	$MC.sd_2$	$MC.sd_3$	$asy.sd_1$	$asy.sd_2$	$asy.sd_3$	$boot.sd_1$	$boot.sd_2$	$boot.sd_3$
						cut-point = (2,4)						
True	0.5000	0.3031	0.8031									
FI	0.4997	0.3037	0.8055	0.0498	0.0338	0.0498	0.0453	0.0294	0.0529	0.0498	0.0348	0.0489
MSI	0.4999	0.3035	0.8046	0.0519	0.0453	0.0554	0.0480	0.0412	0.0578	0.0522	0.0453	0.0542
IPW	0.5003	0.3033	0.8042	0.0617	0.0632	0.0655	0.0616	0.0617	0.0633	0.0624	0.0627	0.0639
SPE	0.5001	0.3034	0.8044	0.0564	0.0588	0.0636	0.0564	0.0573	0.0615	0.0570	0.0579	0.0624
						cut-point = (2,5)						
True	0.5000	0.4682	0.6651									
FI	0.4997	0.4697	0.6684	0.0498	0.0381	0.0617	0.0453	0.0339	0.0608	0.0498	0.0390	0.0608
MSI	0.4999	0.4691	0.6675	0.0519	0.0503	0.0670	0.0480	0.0460	0.0654	0.0522	0.0499	0.0653
IPW	0.5003	0.4687	0.6675	0.0617	0.0688	0.0763	0.0616	0.0668	0.0741	0.0624	0.0676	0.0742
SPE	0.5001	0.4690	0.6674	0.0564	0.0641	0.0735	0.0564	0.0621	0.0706	0.0570	0.0627	0.0715
						cut-point = (2,7)						
True	0.5000	0.7027	0.3349									
FI	0.4997	0.7037	0.3370	0.0498	0.0378	0.0591	0.0453	0.0353	0.0545	0.0498	0.0384	0.0592
MSI	0.4999	0.7037	0.3367	0.0519	0.0482	0.0626	0.0480	0.0451	0.0588	0.0522	0.0476	0.0632
IPW	0.5003	0.7033	0.3371	0.0617	0.0642	0.0709	0.0616	0.0614	0.0715	0.0624	0.0624	0.0721
SPE	0.5001	0.7038	0.3367	0.0564	0.0603	0.0660	0.0564	0.0581	0.0661	0.0570	0.0587	0.0670
						cut-point = (4,5)						
True	0.8031	0.1651	0.6651									
FI	0.8037	0.1660	0.6684	0.0393	0.0277	0.0617	0.0366	0.0236	0.0608	0.0388	0.0282	0.0608
MSI	0.8033	0.1656	0.6675	0.0425	0.0369	0.0670	0.0400	0.0333	0.0654	0.0420	0.0369	0.0653
IPW	0.8033	0.1654	0.6675	0.0486	0.0497	0.0763	0.0469	0.0486	0.0741	0.0475	0.0496	0.0742
SPE	0.8032	0.1656	0.6674	0.0469	0.0460	0.0735	0.0457	0.0454	0.0706	0.0460	0.0458	0.0715
						cut-point = (4,7)						
True	0.8031	0.3996	0.3349									
FI	0.8037	0.4000	0.3370	0.0393	0.0419	0.0591	0.0366	0.0383	0.0545	0.0388	0.0430	0.0592
MSI	0.8033	0.4002	0.3367	0.0425	0.0513	0.0626	0.0400	0.0480	0.0588	0.0420	0.0519	0.0632
IPW	0.8033	0.4000	0.3371	0.0486	0.0639	0.0709	0.0469	0.0643	0.0715	0.0475	0.0652	0.0721
SPE	0.8032	0.4004	0.3367	0.0469	0.0604	0.0660	0.0457	0.0605	0.0661	0.0460	0.0612	0.0670
						cut-point = (5,7)						
True	0.8996	0.2345	0.3349									
FI	0.9000	0.2340	0.3370	0.0271	0.0348	0.0591	0.0255	0.0313	0.0545	0.0269	0.0356	0.0592
MSI	0.8998	0.2346	0.3367	0.0312	0.0441	0.0626	0.0296	0.0407	0.0588	0.0310	0.0441	0.0632
IPW	0.8998	0.2347	0.3371	0.0359	0.0553	0.0709	0.0347	0.0545	0.0715	0.0355	0.0555	0.0721
SPE	0.8998	0.2348	0.3367	0.0351	0.0521	0.0660	0.0347	0.0516	0.0661	0.0348	0.0521	0.0670

TABLE 4. *Simulation results from 5000 replications when the model for the verification process is misspecified (Study 2) and the second value of* Λ *is used. "True" indicates the true parameter value. Sample size = 1000.*

	TCF_1	TCF_2	TCF_3	$MC.sd_1$	$MC.sd_2$	$MC.sd_3$	$asy.sd_1$	$asy.sd_2$	$asy.sd_3$	$boot.sd_1$	$boot.sd_2$	$boot.sd_3$
						cut-point = (2,4)						
True	0.5000	0.3970	0.8970									
FI	0.4998	0.3970	0.8977	0.0267	0.0211	0.0207	0.0231	0.0172	0.0268	0.0268	0.0213	0.0204
MSI	0.4997	0.3970	0.8978	0.0272	0.0240	0.0220	0.0237	0.0202	0.0279	0.0274	0.0239	0.0219
IPW	**0.5983**	**0.3743**	**0.9150**	0.0364	0.0407	0.0257	0.0368	0.0399	0.0284	0.0369	0.0399	0.0258
SPE	0.4996	0.3971	0.8980	0.0314	0.0342	0.0268	0.0314	0.0336	0.0267	0.0315	0.0336	0.0268
						cut-point = (2,5)						
True	0.5000	0.6335	0.7365									
FI	0.4998	0.6340	0.7381	0.0267	0.0218	0.0332	0.0231	0.0198	0.0344	0.0268	0.0219	0.0326
MSI	0.4997	0.6338	0.7381	0.0272	0.0246	0.0344	0.0237	0.0226	0.0356	0.0274	0.0243	0.0338
IPW	**0.5983**	**0.5749**	**0.7965**	0.0364	0.0406	0.0348	0.0368	0.0401	0.0387	0.0369	0.0402	0.0346
SPE	0.4996	0.6338	0.7383	0.0314	0.0341	0.0368	0.0314	0.0335	0.0364	0.0315	0.0336	0.0364
						cut-point = (2,7)						
True	0.5000	0.8682	0.2635									
FI	0.4998	0.8690	0.2639	0.0267	0.0175	0.0295	0.0231	0.0190	0.0280	0.0268	0.0176	0.0290
MSI	0.4997	0.8689	0.2639	0.0272	0.0197	0.0308	0.0237	0.0211	0.0292	0.0274	0.0198	0.0302
IPW	**0.5983**	**0.8307**	**0.3054**	0.0364	0.0342	0.0347	0.0368	0.0341	0.0358	0.0369	0.0343	0.0343
SPE	0.4996	0.8688	0.2639	0.0314	0.0283	0.0316	0.0314	0.0282	0.0308	0.0315	0.0284	0.0310
						cut-point = (4,5)						
True	0.8970	0.2365	0.7365									
FI	0.8975	0.2370	0.7381	0.0159	0.0191	0.0332	0.0153	0.0147	0.0344	0.0162	0.0191	0.0326
MSI	0.8974	0.2368	0.7381	0.0168	0.0215	0.0344	0.0162	0.0175	0.0356	0.0171	0.0213	0.0338
IPW	**0.9216**	**0.2006**	**0.7965**	0.0165	0.0315	0.0348	0.0167	0.0308	0.0387	0.0168	0.0309	0.0346
SPE	0.8974	0.2367	0.7383	0.0189	0.0266	0.0368	0.0191	0.0261	0.0364	0.0191	0.0262	0.0364
						cut-point = (4,7)						
True	0.8970	0.4711	0.2635									
FI	0.8975	0.4721	0.2639	0.0159	0.0276	0.0295	0.0153	0.0247	0.0280	0.0162	0.0276	0.0290
MSI	0.8974	0.4719	0.2639	0.0168	0.0300	0.0308	0.0162	0.0269	0.0292	0.0171	0.0296	0.0302
IPW	**0.9216**	**0.4564**	**0.3054**	0.0165	0.0395	0.0347	0.0167	0.0387	0.0358	0.0168	0.0388	0.0343
SPE	0.8974	0.4717	0.2639	0.0189	0.0339	0.0316	0.0191	0.0331	0.0308	0.0191	0.0333	0.0310
						cut-point = (5,7)						
True	0.9711	0.2347	0.2635									
FI	0.9712	0.2351	0.2639	0.0069	0.0208	0.0295	0.0067	0.0185	0.0280	0.0070	0.0209	0.0290
MSI	0.9712	0.2351	0.2639	0.0083	0.0237	0.0308	0.0080	0.0214	0.0292	0.0084	0.0234	0.0302
IPW	**0.9752**	**0.2558**	**0.3054**	0.0092	0.0319	0.0347	0.0091	0.0315	0.0358	0.0092	0.0316	0.0343
SPE	0.9713	0.2350	0.2639	0.0101	0.0273	0.0316	0.0100	0.0266	0.0308	0.0101	0.0267	0.0310

$$D_3 = \begin{cases} 1 & \text{if } Z_1 + Z_2 > h_2 \\ 0 & \text{otherwise} \end{cases}.$$

Here, h_1 and h_2 are two thresholds. We choose h_1 and h_2 to make $\theta_1 = 0.4$ and $\theta_3 = 0.25$. The continuous test results T and the covariate A are generated to be related to D through Z_1 and Z_2. More precisely,

$$T = 0.5(Z_1 + Z_2) + \varepsilon_1, \qquad A = Z_1 + Z_2 + \varepsilon_2,$$

where ε_1 and ε_2 are two independent normal random variables with mean 0 and the common variance 0.25, independent also from Z_1 and Z_2. The verification status V is simulated by the following logistic model

$$\text{logit}\left\{\Pr(V = 1|T, A)\right\} = 0.1 - 1.53T + A.$$

Under this model, the verification rate is roughly 0.52. We consider the cut points c_1 and c_2 as the pairs $(-1, -0.5), (-1, 0.7), (-1, 1.3), (-0.5, 0.7)$, $(-0.5, 1.3)$ and $(0.7, 1.3)$. In this set–up, we determine the true values of TCF's as

$$\begin{aligned}
\text{TCF}_1(c_1) &= \frac{1}{\Phi(h_1)} \int_{-\infty}^{h_1} \Phi\left(\frac{c_1 - 0.5z}{\sqrt{0.25}}\right) \phi(z)\mathrm{d}z, \\
\text{TCF}_2(c_1, c_2) &= \frac{1}{\Phi(h_2) - \Phi(h_1)} \int_{h_1}^{h_2} \left[\Phi\left(\frac{c_2 - 0.5z}{\sqrt{0.25}}\right) - \Phi\left(\frac{c_1 - 0.5z}{\sqrt{0.25}}\right)\right] \phi(z)\mathrm{d}z, \\
\text{TCF}_3(c_2) &= 1 - \frac{1}{1 - \Phi(h_2)} \int_{h_2}^{\infty} \Phi\left(\frac{c_2 - 0.5z}{\sqrt{0.25}}\right) \phi(z)\mathrm{d}z.
\end{aligned}$$

The aim in this scenario is to evaluate the behavior of the estimators, in particular that of FI, MSI and SPE, when the estimators $\hat{\rho}_{ki}$ are inconsistent, whereas $\hat{\pi}_i$ are consistent. Therefore, $\hat{\rho}_{ki}$ are obtained from a multinomial logistic regression model with $\mathcal{D} = (D_1, D_2, D_3)$ as the response and T as predictor. As the correct process is a multinomial probit process, the chosen model is clearly misspecified. To estimate the conditional verification process π, we use a generalized linear model for V given T and A with logit link. Clearly, this model is correctly specified.

Table 5 shows Monte Carlo means and standard deviations for the estimators of the true class fractions TCF_1, TCF_2 and TCF_3. Moreover, estimated standard deviations (via asymptotic theory) and bootstrap standard deviations are also presented. The results clearly show the effect of misspecification on FI and MSI estimates, despite the high sample size. In particular, in terms of bias, the two methods performs almost alway poorly, with high distortion in some cases (values highlighted in bold). Again, the SPE estimator behaves well due to its doubly robustness property.

4.4. Study 4

We generate data exactly as in Study 3. The aim in this scenario is to evaluate the behavior of FI, MSI, IPW and SPE estimators when the estimates $\hat{\rho}_{ki}$ and

TABLE 5. *Simulation results from 5000 replications when only model for ρ_k is misspecified (Study 3). "True" indicates the true parameter value. Sample size = 1000.*

	TCF_1	TCF_2	TCF_3	$MC.sd_1$	$MC.sd_2$	$MC.sd_3$	$asy.sd_1$	$asy.sd_2$	$asy.sd_3$	$boot.sd_1$	$boot.sd_2$	$boot.sd_3$
						cut-point = (-1,-0.5)						
True	0.1812	0.1070	0.9817									
FI	**0.2144**	**0.1318**	0.9813	0.0230	0.0152	0.0051	0.0243	0.0135	0.0200	0.0230	0.0150	0.0052
MSI	**0.2172**	**0.1328**	0.9800	0.0237	0.0182	0.0074	0.0250	0.0166	0.0207	0.0237	0.0179	0.0075
IPW	0.1819	0.1072	0.9817	0.0258	0.0197	0.0091	0.0258	0.0194	0.0135	0.0260	0.0196	0.0092
SPE	0.1818	0.1073	0.9816	0.0208	0.0206	0.0093	0.0207	0.0202	0.0090	0.0208	0.0204	0.0094
						cut-point = (-1,0.7)						
True	0.1812	0.8609	0.4469									
FI	**0.2144**	**0.8879**	**0.4010**	0.0230	0.0149	0.0284	0.0243	0.0153	0.0242	0.0230	0.0146	0.0284
MSI	**0.2172**	**0.8931**	**0.4035**	0.0237	0.0165	0.0292	0.0250	0.0172	0.0251	0.0237	0.0165	0.0292
IPW	0.1819	0.8606	0.4462	0.0258	0.0350	0.0437	0.0258	0.0342	0.0447	0.0260	0.0348	0.0437
SPE	0.1818	0.8608	0.4462	0.0208	0.0311	0.0455	0.0207	0.0305	0.0449	0.0208	0.0310	0.0482
						cut-point = (-1,1.3)						
True	0.1812	0.9732	0.1171									
FI	**0.2144**	0.9672	**0.0949**	0.0230	0.0063	0.0161	0.0243	0.0099	0.0104	0.0230	0.0062	0.0161
MSI	**0.2172**	0.9708	**0.0960**	0.0237	0.0079	0.0164	0.0250	0.0110	0.0109	0.0237	0.0078	0.0164
IPW	0.1819	0.9734	0.1164	0.0258	0.0167	0.0358	0.0258	0.0130	0.0347	0.0260	0.0160	0.0354
SPE	0.1818	0.9734	0.1169	0.0208	0.0158	0.0281	0.0207	0.0128	0.0263	0.0208	0.0151	0.0333
						cut-point = (-0.5,0.7)						
True	0.4796	0.7539	0.4469									
FI	**0.5497**	0.7561	**0.4010**	0.0302	0.0196	0.0284	0.0284	0.0183	0.0242	0.0301	0.0192	0.0284
MSI	**0.5502**	0.7603	**0.4035**	0.0312	0.0220	0.0292	0.0295	0.0211	0.0251	0.0310	0.0219	0.0292
IPW	0.4801	0.7534	0.4462	0.0390	0.0373	0.0437	0.0384	0.0371	0.0447	0.0387	0.0374	0.0437
SPE	0.4801	0.7535	0.4462	0.0327	0.0344	0.0455	0.0322	0.0339	0.0449	0.0324	0.0343	0.0482
						cut-point = (-0.5,1.3)						
True	0.4796	0.8661	0.1171									
FI	**0.5497**	**0.8354**	**0.0949**	0.0302	0.0189	0.0161	0.0284	0.0185	0.0104	0.0301	0.0186	0.0161
MSI	**0.5502**	**0.8380**	**0.0960**	0.0312	0.0207	0.0164	0.0295	0.0204	0.0109	0.0310	0.0204	0.0164
IPW	0.4801	0.8661	0.1164	0.0390	0.0248	0.0358	0.0384	0.0238	0.0347	0.0387	0.0245	0.0354
SPE	0.4801	0.8660	0.1169	0.0327	0.0250	0.0281	0.0322	0.0239	0.0263	0.0324	0.0245	0.0333
						cut-point = (0.7,1.3)						
True	0.9836	0.1122	0.1171									
FI	**0.9933**	**0.0793**	**0.0949**	0.0023	0.0133	0.0161	0.0021	0.0119	0.0104	0.0023	0.0131	0.0161
MSI	**0.9930**	**0.0777**	**0.0960**	0.0038	0.0145	0.0164	0.0032	0.0135	0.0109	0.0038	0.0145	0.0164
IPW	0.9839	0.1128	0.1164	0.0183	0.0324	0.0358	0.0122	0.0319	0.0347	0.0173	0.0325	0.0354
SPE	0.9839	0.1125	0.1169	0.0180	0.0283	0.0281	0.0122	0.0280	0.0263	0.0170	0.0285	0.0333

$\hat{\pi}_i$ are inconsistent. Therefore, $\hat{\rho}_{ki}$ are obtained from a multinomial logistic regression model with $\mathcal{D} = (D_1, D_2, D_3)$ as the response and T as predictor. This model is misspecified. To estimate the conditional verification disease π, we use a generalized linear model for V given T and $A^{2/3}$ with logit link. Clearly, this model is misspecified.

Table 6 shows Monte Carlo means and standard deviations for the estimators of the true class fractions TCF_1, TCF_2 and TCF_3. Moreover, estimated standard deviations (via asymptotic theory) and bootstrap standard deviations are also presented. The results clearly show that when both the disease and verification models are misspecified, all estimators may behave poorly, with high distortion in some cases (values highlighted in bold).

5. Two illustrations

To illustrate the application of the proposed methods, in this section we consider two quite distinct real data examples, both dealing with epithelial ovarian cancer (EOC). In the first illustration, we consider diagnosis of EOC in one of three classes i.e., benign disease, early stage and late stage cancer on the basis of a well known tumor marker, i.e., CA125. We make use of a publicly available dataset in which the disease status is known for all subjects. Then, we simulate a verification process and apply our estimators. This allows to compare results obtained in the complete data case with those obtained in the incomplete data case after correcting for verification bias. In the second illustration, we focus on prediction of patients' response to chemotherapy, classified as sensitive, partially sensitive and resistant. Data are available for late stage EOC patients. In this second example, the response is missing for about 25% of the subjects involved in the study.

5.1. Diagnosis of EOC

We use data from the Pre-PLCO Phase II Dataset from the SPORE/Early Detection Network/Prostate, Lung, Colon, and Ovarian Cancer Ovarian Validation Study. The study protocol and data are publicly available at the address[1], along with descriptions of the study aims and analytic methods. In particular, we consider the following three classes of EOC, i.e., benign disease, early stage (I and II) and late stage (III and IV) cancer, and 2 of the 59 available biomarkers, i.e. CA125 and CA153, measured at Harvard laboratories. In detail, we use CA125 as the test T s and CA153 as a covariate. Reasons for using CA153 as a covariate come from the medical literature that suggests that the concomitant measurement of CA153 with CA125 could be advantageous in the pre-operative discrimination of benign and malignant ovarian tumors. In addition, age of patients is also considered. Here, we have 134 patients with benign disease, 67 early stage samples and 77 late stage samples.

[1] http://edrn.nci.nih.gov/protocols/119-spore-edrn-pre-plco-ovarian-phase-ii-validation

TABLE 6. *Simulation results from 5000 replications when both models for ρ_k and π are misspecified (Study 4). "True" indicates the true parameter value. Sample size = 1000.*

	TCF_1	TCF_2	TCF_3	$MC.sd_1$	$MC.sd_2$	$MC.sd_3$	$asy.sd_1$	$asy.sd_2$	$asy.sd_3$	$boot.sd_1$	$boot.sd_2$	$boot.sd_3$
						cut-point = (-1,-0.5)						
True	0.1812	0.1070	0.9817									
FI	**0.2143**	**0.1320**	0.9814	0.0231	0.0149	0.0051	0.0243	0.0135	0.0200	0.0230	0.0150	0.0052
MSI	**0.2170**	**0.1330**	0.9801	0.0238	0.0179	0.0074	0.0250	0.0166	0.0207	0.0237	0.0179	0.0075
IPW	**0.2185**	**0.1339**	0.9792	0.0284	0.0234	0.0102	0.0282	0.0232	0.0105	0.0283	0.0233	0.0102
SPE	**0.2183**	**0.1339**	0.9792	0.0247	0.0220	0.0101	0.0245	0.0219	0.0098	0.0246	0.0219	0.0102
						cut-point = (-1,0.7)						
True	0.1812	0.8609	0.4469									
FI	**0.2143**	**0.8887**	**0.4002**	0.0231	0.0143	0.0285	0.0243	0.0153	0.0242	0.0230	0.0146	0.0285
MSI	**0.2170**	**0.8940**	**0.4029**	0.0238	0.0164	0.0290	0.0250	0.0171	0.0251	0.0237	0.0165	0.0292
IPW	**0.2185**	**0.8994**	**0.4078**	0.0284	0.0237	0.0397	0.0282	0.0232	0.0410	0.0283	0.0234	0.0397
SPE	**0.2183**	**0.8998**	**0.4071**	0.0247	0.0223	0.0323	0.0245	0.0219	0.0325	0.0246	0.0220	0.0326
						cut-point = (-1,1.3)						
True	0.1812	0.9732	0.1171									
FI	**0.2143**	0.9675	**0.0947**	0.0231	0.0061	0.0160	0.0243	0.0099	0.0104	0.0230	0.0062	0.0161
MSI	**0.2170**	0.9711	**0.0958**	0.0238	0.0078	0.0163	0.0250	0.0110	0.0108	0.0237	0.0078	0.0164
IPW	**0.2185**	0.9742	**0.0977**	0.0284	0.0112	0.0269	0.0282	0.0107	0.0270	0.0283	0.0111	0.0273
SPE	**0.2183**	0.9742	**0.0978**	0.0247	0.0110	0.0174	0.0245	0.0105	0.0175	0.0246	0.0108	0.0177
						cut-point = (-0.5,0.7)						
True	0.4796	0.7539	0.4469									
FI	**0.5510**	0.7567	**0.4002**	0.0306	0.0190	0.0285	0.0285	0.0183	0.0242	0.0301	0.0192	0.0285
MSI	**0.5514**	0.7610	**0.4029**	0.0316	0.0219	0.0290	0.0295	0.0211	0.0251	0.0310	0.0219	0.0292
IPW	**0.5509**	0.7655	**0.4078**	0.0360	0.0313	0.0397	0.0357	0.0310	0.0410	0.0358	0.0311	0.0397
SPE	**0.5509**	0.7659	**0.4071**	0.0336	0.0286	0.0323	0.0329	0.0286	0.0325	0.0329	0.0287	0.0326
						cut-point = (-0.5,1.3)						
True	0.4796	0.8661	0.1171									
FI	**0.5510**	**0.8355**	**0.0947**	0.0306	0.0186	0.0160	0.0285	0.0186	0.0104	0.0301	0.0186	0.0161
MSI	**0.5514**	**0.8380**	**0.0958**	0.0316	0.0205	0.0163	0.0295	0.0204	0.0108	0.0310	0.0204	0.0164
IPW	**0.5509**	**0.8403**	**0.0977**	0.0360	0.0255	0.0269	0.0357	0.0251	0.0270	0.0358	0.0252	0.0273
SPE	**0.5509**	**0.8403**	**0.0978**	0.0336	0.0240	0.0174	0.0329	0.0237	0.0175	0.0329	0.0238	0.0177
						cut-point = (0.7,1.3)						
True	0.9836	0.1122	0.1171									
FI	**0.9934**	**0.0788**	**0.0947**	0.0022	0.0129	0.0160	0.0021	0.0119	0.0104	0.0023	0.0131	0.0161
MSI	**0.9930**	**0.0771**	**0.0958**	0.0038	0.0145	0.0163	0.0032	0.0134	0.0108	0.0037	0.0145	0.0164
IPW	**0.9925**	**0.0748**	**0.0977**	0.0075	0.0213	0.0269	0.0057	0.0208	0.0270	0.0073	0.0211	0.0273
SPE	**0.9925**	**0.0744**	**0.0978**	0.0074	0.0201	0.0174	0.0058	0.0196	0.0175	0.0073	0.0198	0.0177

To mimic verification bias, a subset of the complete dataset is constructed using the test T and the vector $A = (A_1, A_2)^\top$ of the two covariates, namely the marker CA153 (A_1) and age (A_2). In this subset, T and A are known for all samples, but the true status (benign, early stage or late stage) is available only for some samples, that we select according to the following mechanism. We select all samples having a value for both T and A above their respective medians, i.e. 0.87 and (45,0.30); as for the others, we apply the following selection process

$$\Pr(V = 1) = 0.05 + \delta_1 \mathrm{I}(T > 0.87) + \delta_2 \mathrm{I}(A_1 > 0.30) + \delta_3 \mathrm{I}(A_2 > 45),$$

with $\delta_1 = 0.35$, $\delta_2 = 0.25$ and $\delta_3 = 0.35$, leading to a marginal probability of selection equal to 0.634. With such a choice, the verification probability is equal to about 0.65 for subjects with $T > 0.87$, $A_1 > 0.30$ and $A_2 < 45$; 0.75 for subjects with $T > 0.87$, $A_1 < 0.30$ and $A_2 > 45$; 0.65 for subjects with $T < 0.87$, $A_1 > 0.30$ and $A_2 > 45$; 0.4 for subjects with $T > 0.87$, $A_1 < 0.30$ and $A_2 < 45$; 0.3 for subjects with $T < 0.87$, $A_1 > 0.30$ and $A_2 < 45$; 0.4 for subjects with $T < 0.87$, $A_1 < 0.30$ and $A_2 > 45$; 0.05 otherwise.

To apply FI, MSI and SPE estimators, we employ a multinomial logistic model to estimate $\rho_{ki} = \Pr(D_{ki} = 1|T_i, A_{1i}, A_{2i})$, where $D_k = 1$, $k = 1, 2, 3$ refers to benign, early and late, respectively. On the other hand, SPE and IPW methods require estimates of $\pi_i = \Pr(V_i = 1|T_i, A_{1i}, A_{2i})$. For estimating such quantities, we make use, firstly, of a correctly specified model, i.e., a linear threshold regression model and, then, of a misspecified model, i.e., a logistic model.

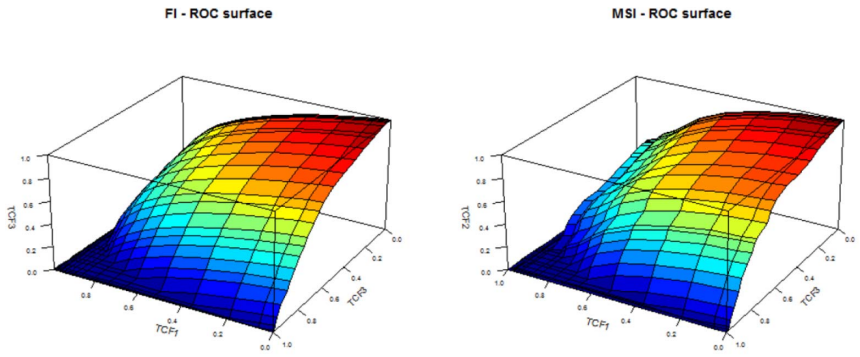
The estimated ROC surfaces for the test T (CA125) obtained by applying the proposed methods are shown in Figure 1.

For the sake of comparison, we also produced the estimate of the ROC surface with full data (Full estimate), reported in Appendix D, Figure 3. In Appendix D, Figure 4 and Figure 5, we also give the projections of the bias–corrected estimated ROC surfaces to the planes defined by TCF_1 versus TCF_2, TCF_1 versus TCF_3 and TCF_2 versus TCF_3, i.e., the ROC curves between classes 1 and 2, classes 1 and 3, classes 2 and 3. Such plots are obtained by setting $\mathrm{TCF}_3 = 0$, $\mathrm{TCF}_2 = 0$ and $\mathrm{TCF}_1 = 0$, respectively. For example, the estimated ROC curves between classes 1 and 2 are defined as the set of points

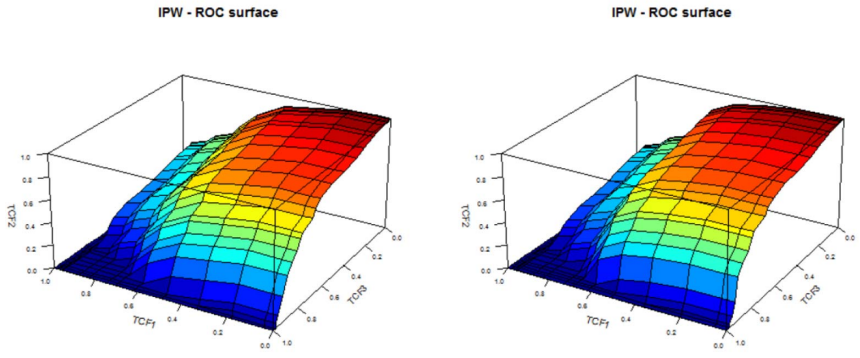
$$\left\{ (\widehat{\mathrm{TCF}}_{1,*}(c_1), \widehat{\mathrm{TCF}}_{2,*}(c_1, +\infty)), c_1 \in \mathbb{R} \right\},$$

that is, we ignore the cut point c_2. This is a construction equivalent to the most popular representation of an estimated ROC curve, which usually depicts $\widehat{\mathrm{TCF}}_2$ versus $1 - \widehat{\mathrm{TCF}}_1$.

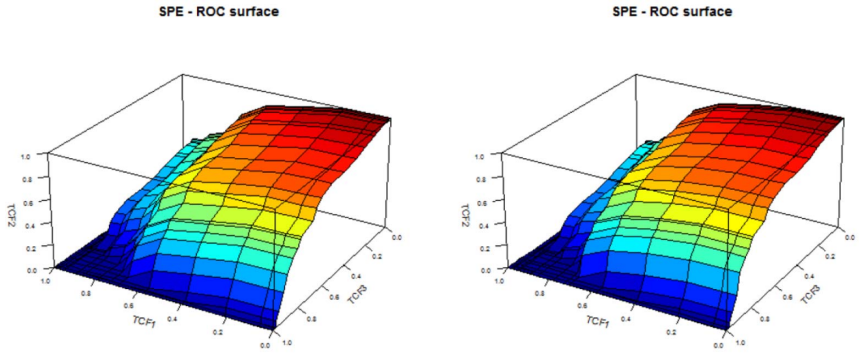
Compared with the Full estimate, all the bias-corrected methods discussed in the paper seem to behave well, yielding reasonable estimates of the ROC surface and the ROC curves. Moreover, Table 7 shows the VUS estimates obtained with the FI, MSI, IPW and SPE estimators (both for the correctly specified and misspecified model for the verification process), along with approximated 95% confidence intervals. Inspection of the table highlights that estimators with better performance are, overall, IPW and SPE. This might be an indication that



(a) FI (b) MSI



(c) IPW–threshold model (d) IPW–logit model



(e) SPE–threshold model (f) SPE–logit model

FIG 1. *Bias–corrected estimated ROC surfaces for CA125, assessing the classification into three class of EOC: benign disease, early stage (I and II) and late stage (III and IV). For the SPE and IPW approaches, results for the correctly specified and misspecified model for the verification process are given.*

TABLE 7

Bias–corrected (and Full) estimated VUS for the marker CA125, assessing the classification into three classes of EOC: benign disease, early stage (I and II) and late stage (III and IV). For the SPE and IPW approaches, results for the correctly specified and misspecified model for the verification process are given.

	VUS Estimate	Asy.sd	Boot.sd	95% C.I. (with Asy.sd)
Full	0.5663			
FI	0.5150	0.0404	0.0417	(0.4357, 0.5942)
MSI	0.5183	0.0415	0.0431	(0.4368, 0.5997)
IPW.logit	0.5500	0.0416	0.0471	(0.4685, 0.6314)
SPE.logit	0.5581	0.0443	0.0463	(0.4712, 0.6450)
IPW.thres	0.5353	0.0393	0.0457	(0.4583, 0.6123)
SPE.thres	0.5470	0.0440	0.0438	(0.4608, 0.6331)

the multinomial logistic model chosen for the disease process might not be fully adequate in this case.

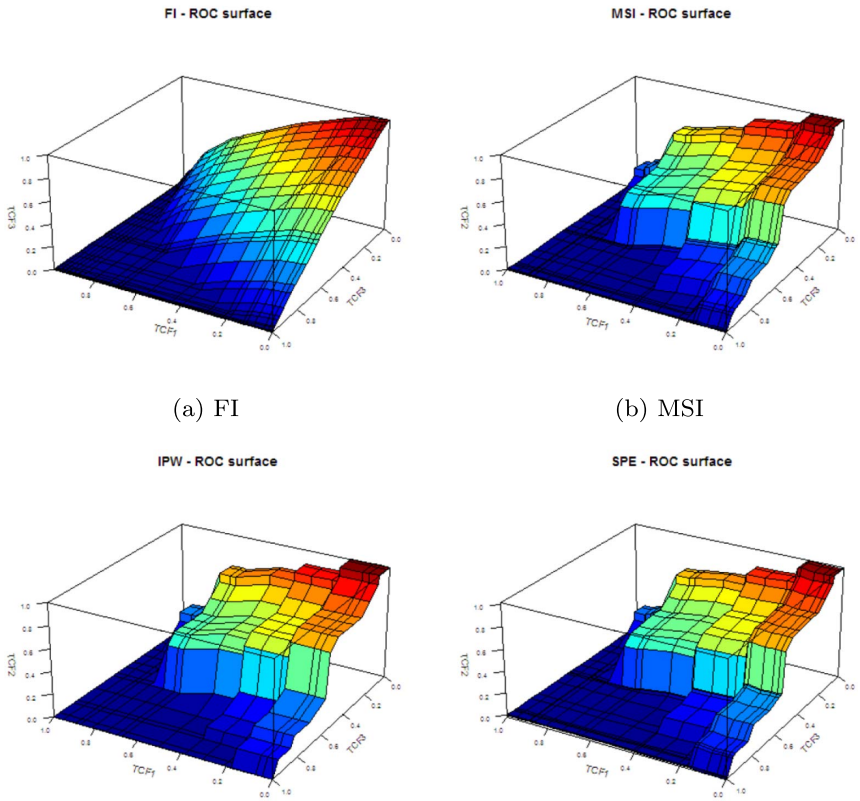
5.2. Prediction of response to chemotherapy

A major challenge in advanced-stage EOC is prediction of response to platinum chemotherapy on the basis of markers measured at molecular level. Indeed, several genomic profiling studies have shown that gene expressions relate with different aspects of ovarian cancer (tumor subtype, stage, grade, prognosis, and therapy resistance), although the measured association is usually very low. Here, we consider a cohort of 99 snap-frozen tumor biopsies taken from a frozen tissue bank, located at the Department of Oncology, IRCCS-Mario Negri Institute, Milano, Italy. Biopsies were collected from late stage (III and IV) cancer patients who underwent surgery at the Obstetrics and Gynaecology Department, San Gerardo Hospital, Monza, Italy between September 1992 and March 2010. For 75 of the 99 subjects, the three-class response to platinum therapy is available, being 31 patients sensitive, 11 partially sensitive and 33 resistant. For all the subjects, we consider as test predictive of the response to therapy the marker (T) resulting as a given linear combination of the logarithm of the expression levels of six genes, i.e., Entrez Gene ID: 23513, 7284, 128408, 56996, 2969, 6170. As a covariate, we consider age at onset of patients.

The estimated ROC surfaces for T obtained by applying the proposed methods are shown in Figure 2. FI, MSI, IPW and SPE estimators are based on the multinomial logistic model for the disease process and/or the logistic model for the verification process. Table 8 shows the corresponding VUS estimates, along with the naïve estimate. The table also gives the estimated standard deviations (via asymptotic theory), bootstrap standard deviations and approximated 95% confidence intervals. Despite the limited sample size, the results show that T has some ability to predict response to therapy for late stage EOC patients.

6. Conclusions

This paper proposed several verification bias-corrected estimators of the ROC surface (and the VUS) of a continuous diagnostic test. These estimators, which



(a) FI (b) MSI

(c) IPW (d) SPE

FIG 2. *Bias–corrected estimated ROC surfaces for the test T predicting the response to therapy of late stage EOC patients.*

TABLE 8
Bias–corrected (and Naïve) estimated VUS for the test T predicting the response to therapy of late stage EOC patients.

	VUS Estimate	Asy.sd	Boot.sd	95% C.I. (with Asy.sd)
Naïve	0.3452			
FI	0.3005	0.0512	0.0538	(0.2002, 0.4009)
MSI	0.3197	0.0629	0.0656	(0.1963, 0.4430)
IPW	0.3231	0.0654	0.0755	(0.1949, 0.4512)
SPE	0.3110	0.0675	0.0704	(0.1787, 0.4433)

can be considered an extension to the three-class case of estimators in Alonzo and Pepe [3], are partially parametric in that they require the choice of a parametric model for the estimation of the disease process, or of the verification process, or of both processes. In some cases, wrong specifications of such models can visibly affect the produced estimates, as highlighted also by our results in

the simulation studies. In fact, FI and MSI estimators are inconsistent if the model for the disease process is misspecified. On the contrary, the IPW estimator is inconsistent if the model for the verification process is misspecified. Thanks to the property of doubly robustness, inconsistency of SPE estimators occurs only if both models for the two processes are misspecified.

Generally speaking, to avoid misspecification problems, one possibility could be to resort on fully nonparametric estimators. It is worth noting, however, that misspecification could be also caused by a wrong assumption on the missingness mechanism. Indeed, the verification process could somehow depend, in addition to the test T and covariates, also on the presumed disease status, a possibility not foreseen by the MAR assumption. In such situations, we face a nonignorable missingness mechanism and such setup is not dealt with by the proposed estimators. Both topics, i.e., finding nonparametric estimators of ROC surfaces under MAR assumption and dealing with nonignorable missingness, are the focus of our current work.

Appendix A: Asymptotic distribution results

In this section, we discuss validity of conditions (C1), (C2) and (C3) for the proposed estimators. The discussion covers first the elements of the estimating functions corresponding to the parameter τ. Then, we pass to the elements of the estimating functions corresponding to the parameters $\theta_1, \theta_2, \theta_{11}, \beta_{12}, \beta_{22}, \beta_{23}$, specializing the discussion to the various methods. Finally, we give the explicit form of the variance-covariance matrix in Theorem 1. Recall that α_0 denotes the true value of α.

Parameter τ. We noted in Section 3 that estimators FI, MSI and SPE require a multinomial logistic or probit regression model to estimate the disease probabilities $\rho_{ki} = \Pr(D_{ki} = 1|T_i, A_i)$ with $k = 1, 2, 3$. In the following, we adopt the multinomial logistic model, but arguments similar to those given below hold also for the multinomial probit model, despite the rather more complex algebra (see Daganzo [5], Chapter 3, as a general reference).

The estimating function for the nuisance parameter $\tau \equiv \tau_\rho = (\tau_{\rho_1}^\top, \tau_{\rho_2}^\top)^\top$,

$$G^{\tau_\rho}(\alpha) = \left(G^{\tau_{\rho_1}}(\alpha)^\top, G^{\tau_{\rho_2}}(\alpha)^\top\right)^\top \equiv \left(\left(\sum_{i=1}^n g_i^{\tau_{\rho_1}}(\alpha)\right)^\top, \left(\sum_{i=1}^n g_i^{\tau_{\rho_2}}(\alpha)\right)^\top\right)^\top,$$

is obtained as the first derivative of the log likelihood function. With the multinomial logistic model, we get

$$G^{\tau_\rho}(\alpha) = \left(\left(\sum_{i=1}^n V_i U_i (D_{1i} - \rho_{1i})\right)^\top, \left(\sum_{i=1}^n V_i U_i (D_{2i} - \rho_{2i})\right)^\top\right)^\top,$$

where $U_i = (1, T_i, A_i)^\top$. Under assumption (A2), condition (C1) trivially holds.

Moreover, we get

$$\begin{aligned}
\frac{\partial}{\partial \tau_{\rho_1}^{\top}} g_i^{\tau_{\rho_1}}(\alpha) &= -V_i U_i U_i^{\top} \rho_{1i}(1-\rho_{1i}), \quad \frac{\partial}{\partial \tau_{\rho_2}^{\top}} g_i^{\tau_{\rho_1}}(\alpha) = V_i U_i U_i^{\top} \rho_{1i}\rho_{2i}, \\
\frac{\partial}{\partial \tau_{\rho_2}^{\top}} g_i^{\tau_{\rho_2}}(\alpha) &= -V_i U_i U_i^{\top} \rho_{2i}(1-\rho_{2i}), \quad \frac{\partial}{\partial \tau_{\rho_1}^{\top}} g_i^{\tau_{\rho_2}}(\alpha) = V_i U_i U_i^{\top} \rho_{1i}\rho_{2i},
\end{aligned} \tag{A.1}$$

and $\frac{\partial}{\partial \theta_s} g_i^{\tau_\rho}(\alpha) = 0$, $\frac{\partial}{\partial \beta_{jk}} g_i^{\tau_\rho}(\alpha) = 0$ for each s, j, k. The second–order partial derivatives can be easily derived. Hence, for $G^{\tau_\rho}(\alpha)$, condition (C2) holds and, by assumption (A3)–(A5) condition (C3) also holds.

The IPW and SPE estimators require estimates of $\pi_i = \Pr(V_i = 1|T_i, A_i)$. With T and A as covariates, we can use the logistic or probit models to this end. In these cases, conditions (C1)–(C3) are satisfied by the score functions

$$G^{\tau_\pi}(\alpha) = \sum_{i=1}^{n} g_i^{\tau_\pi}(\alpha) = \sum_{i=1}^{n} U_i(V_i - \pi_i)$$

or

$$G^{\tau_\pi}(\alpha) = \sum_{i=1}^{n} g_i^{\tau_\pi}(\alpha) = \sum_{i=1}^{n} \left[\frac{V_i U_i \phi(U_i^{\top}\tau_\pi)}{\Phi(U_i^{\top}\tau_\pi)} - (1-V_i)\frac{U_i \phi(U_i^{\top}\tau_\pi)}{1-\Phi(U_i^{\top}\tau_\pi)} \right],$$

respectively, where $\phi(\cdot)$ and $\Phi(\cdot)$ are the density function and the cumulative distribution function of the standard normal random variable, respectively. Recall that τ_π is the component of nuisance parameter τ corresponding the model for estimating π. The first-order derivatives are

$$\frac{\partial}{\partial \tau_\pi^{\top}} g_i^{\tau_\pi}(\alpha) = -U_i U_i^{\top} \pi_i(1-\pi_i),$$

or

$$\begin{aligned}
\frac{\partial}{\partial \tau_\pi^{\top}} g_i^{\tau_\pi}(\alpha) &= -\frac{V_i U_i U_i^{\top} \phi(U_i^{\top}\tau_\pi)\left[-U_i^{\top}\tau_\pi \Phi(U_i^{\top}\tau_\pi) - \phi(U_i^{\top}\tau_\pi)\right]}{\Phi^2(U_i^{\top}\tau_\pi)} \\
&\quad - (1-V_i)\frac{U_i U_i^{\top} \phi(U_i^{\top}\tau_\pi)\left[U_i^{\top}\tau_\pi(\Phi(U_i^{\top}\tau_\pi)-1) + \phi(U_i^{\top}\tau_\pi)\right]}{\left[1-\Phi(U_i^{\top}\tau_\pi)\right]^2}.
\end{aligned}$$

FI and MSI estimators. According to equations (3.4), (3.5) and (3.6), the estimating functions $G_*^{\theta_s}(\alpha)$ for FI and MSI estimators can be presented in the form

$$G_{\mathrm{IE}}^{\theta_s}(\alpha) \equiv \sum_{i=1}^{n} g_{i,\mathrm{IE}}^{\theta_s}(\alpha) = \sum_{i=1}^{n} \left\{ V_i \left[mD_{si} - \theta_s + (1-m)\rho_{si} \right] + (1-V_i)(\rho_{si} - \theta_s) \right\},$$

with $s = 1, 2$. Similarly,

$$G_{\mathrm{IE}}^{\beta_{jk}}(\alpha) \equiv \sum_{i=1}^{n} g_{i,\mathrm{IE}}^{\beta_{jk}}(\alpha)$$

$$= \sum_{i=1}^{n} \Big\{ V_i \left[m\mathrm{I}(T_i \geq c_j) D_{ki} - \beta_{jk} + (1-m)\mathrm{I}(T_i \geq c_j)\rho_{ki} \right] + (1-V_i)\left[\mathrm{I}(T_i \geq c_j)\rho_{ki} - \beta_{jk} \right] \Big\},$$

for $j = 1, 2$ and $k = 1, 2, 3$. Here, the notation IE means "imputation estimator". The estimating function corresponds to the FI estimator if $m = 0$, to the MSI estimator if $m = 1$. Using the conditional expectation and the assumption (A1), $\mathbb{E}\left[g_{i,\mathrm{IE}}^{\theta_s}(\alpha_0) \right]$ equals

$$\begin{aligned}
&\mathbb{E}_{D_s,T_i,A_i}\left[\mathbb{E}\left[g_{i,\mathrm{IE}}^{\theta_s}(\alpha_0) | T_i, A_i \right] \right] \\
&= \mathbb{E}_{D_{si},T_i,A_i}\left[\mathbb{E}\left[\{ V_i \left[mD_{si} - \theta_{s0} + (1-m)\rho_{si} \right] + (1-V_i)\left[\rho_{si} - \theta_{s0} \right] \} | T_i, A_i \right] \right] \\
&= \mathbb{E}_{D_{si},T_i,A_i}\left[\pi_i \left[m\mathbb{E}\left[D_{si} | T_i, A_i \right] - \theta_{s0} + (1-m)\rho_{si} \right] + (1-\pi_i)(\rho_{si} - \theta_{s0}) \right] \\
&= \mathbb{E}_{D_{si},T_i,A_i}\left[\pi_i \left[m\rho_{si} - \theta_{s0} + (1-m)\rho_{si} \right] + (1-\pi_i)(\rho_{si} - \theta_{s0}) \right] \\
&= \mathbb{E}_{D_{si},T_i,A_i}\left[\pi_i(\rho_{si} - \theta_{s0}) + (1-\pi_i)(\rho_{si} - \theta_{s0}) \right] \\
&= \mathbb{E}_{D_{si},T_i,A_i}\left[\rho_{si} - \theta_{s0} \right] = 0.
\end{aligned}$$

Similarly, we compute the expected value of the estimating function components $g_{i,\mathrm{IE}}^{\beta_{jk}}(\alpha_0)$ as follows

$$\begin{aligned}
&\mathbb{E}_{D_k,T_i,A_i}\left[\mathbb{E}\left[g_{i,\mathrm{IE}}^{\beta_{jk}}(\alpha_0) | T_i, A_i \right] \right] \\
&= \mathbb{E}_{D_{ki},T_i,A_i}\Bigg[\mathbb{E}\Bigg[\Big\{ V_i \left[m\mathrm{I}(T_i \geq c_j) D_{ki} - \beta_{jk0} + (1-m)\mathrm{I}(T_i \geq c_j)\rho_{ki} \right] \\
&+ \ (1-V_i)\left[\mathrm{I}(T_i \geq c_j)\rho_{ki} - \beta_{jk0} \right] \Big\} \Big| T_i, A_i \Bigg] \Bigg] \\
&= \mathbb{E}_{D_{ki},T_i,A_i}\Big[\pi_i \left[m\mathrm{I}(T_i \geq c_j)\rho_{ki} - \beta_{jk0} + (1-m)\mathrm{I}(T_i \geq c_j)\rho_{ki} \right] \\
&+ \ (1-\pi_i)(\mathrm{I}(T_i \geq c_j)\rho_{ki} - \beta_{jk0}) \Big] \\
&= \mathbb{E}_{D_{ki},T_i,A_i}\left[\pi_i(\mathrm{I}(T_i \geq c_j)\rho_{ki} - \beta_{jk0}) + (1-\pi_i)(\mathrm{I}(T_i \geq c_j)\rho_{ki} - \beta_{jk0}) \right] \\
&= \mathbb{E}_{D_{ki},T_i,A_i}\left[\mathrm{I}(T_i \geq c_j)\rho_{ki} - \beta_{jk0} \right] = 0.
\end{aligned}$$

Hence, under assumption (A2), condition (C1) holds for $G_{\mathrm{IE}}^{\theta_s}(\alpha)$ and $G_{\mathrm{IE}}^{\beta_{jk}}(\alpha)$.

We now verify conditions (C2) and (C3). The partial derivative of $G_{\mathrm{IE}}^{\theta_s}(\alpha)$ with respect to β_{jk} equals 0 for all j, k. Moreover,

$$\begin{aligned}
\frac{\partial}{\partial \theta_{s'}} G_{\mathrm{IE}}^{\theta_s}(\alpha) &= \sum_{i=1}^{n} \frac{\partial}{\partial \theta_{s'}} \left\{ V_i \left[mD_{si} - \theta_s + (1-m)\rho_{si} \right] + (1-V_i)\left[\rho_{si} - \theta_s \right] \right\} \\
&= \sum_{i=1}^{n} \mathrm{I}(s' = s)\left\{ -V_i - (1-V_i) \right\} = -n\mathrm{I}(s' = s)
\end{aligned}$$

and

$$\frac{\partial}{\partial \tau_\rho} G_{\mathrm{IE}}^{\theta_s}(\alpha) = \left(\frac{\partial}{\partial \tau_{\rho_1}} G_{\mathrm{IE}}^{\theta_s}(\alpha)^\top, \frac{\partial}{\partial \tau_{\rho_2}} G_{\mathrm{IE}}^{\theta_s}(\alpha)^\top \right)^\top.$$

For each $l = 1, 2$ and $s = 1, 2$, we have

$$\frac{\partial}{\partial \tau_{\rho_l}} G_{\mathrm{IE}}^{\theta_s}(\alpha) = \sum_{i=1}^{n} (1 - mV_i) \frac{\partial}{\partial \tau_{\rho_l}} \rho_{si}.$$

Recall that, under the multinomial logistic model,

$$\rho_{si} = \frac{e^{U_i^\top \tau_s}}{1 + e^{U_i^\top \tau_{\rho_1}} + e^{U_i^\top \tau_{\rho_2}}}, \qquad s = 1, 2.$$

Thus, we obtain

$$\begin{aligned} \tfrac{\partial}{\partial \tau_{\rho_1}} \rho_{1i} &= U_i \rho_{1i}(1 - \rho_{1i}), \quad \tfrac{\partial}{\partial \tau_{\rho_2}} \rho_{1i} = -U_i \rho_{1i} \rho_{2i}, \\ \tfrac{\partial}{\partial \tau_{\rho_2}} \rho_{2i} &= U_i \rho_{2i}(1 - \rho_{2i}), \quad \tfrac{\partial}{\partial \tau_{\rho_1}} \rho_{2i} = -U_i \rho_{1i} \rho_{2i}. \end{aligned} \tag{A.2}$$

The derivatives of $G_{\mathrm{IE}}^{\beta_{jk}}(\alpha)$ are

$$\frac{\partial}{\partial \theta_s} G_{\mathrm{IE}}^{\beta_{jk}}(\alpha) = 0, \qquad \frac{\partial}{\partial \beta_{j'k'}} G_{\mathrm{IE}}^{\beta_{jk}}(\alpha) = -n \mathrm{I}(j'k' = jk)$$

and

$$\frac{\partial}{\partial \tau_{\rho_l}} G_{\mathrm{IE}}^{\beta_{jk}}(\alpha) = \sum_{i=1}^{n} (1 - mV_i) \mathrm{I}(T_i \geq c_j) \frac{\partial}{\partial \tau_{\rho_l}} \rho_{ki},$$

where $\frac{\partial}{\partial \tau_{\rho_l}} \rho_{si}$ is in (A.2). Hence, we have the explicit form of the partial derivatives of both $G_{\mathrm{IE}}^{\theta_s}(\alpha)$ and $G_{\mathrm{IE}}^{\beta_{jk}}(\alpha)$. The only not null elements of the second–order partial derivative of $G_{\mathrm{IE}}^{\theta_s}(\alpha)$ and $G_{\mathrm{IE}}^{\beta_{jk}}(\alpha)$ are those corresponding to the matrices $\frac{\partial^2}{\partial \tau \partial \tau^\top} G_{\mathrm{IE}}^{\theta_s}(\alpha)$ and $\frac{\partial^2}{\partial \tau \partial \tau^\top} G_{\mathrm{IE}}^{\beta_{jk}}(\alpha)$. These elements involve the derivatives with respect to τ of quantities in (A.2). It follows that conditions (C2) and (C3) hold for $G_{\mathrm{IE}}^{\theta_s}(\alpha)$ and $G_{\mathrm{IE}}^{\beta_{jk}}(\alpha)$ for each s, j, k.

IPW estimator. Recall that the estimating function for θ_s is

$$G_{\mathrm{IPW}}^{\theta_s}(\alpha) = \sum_{i=1}^{n} g_{i,\mathrm{IPW}}^{\theta_s}(\alpha) = \sum_{i=1}^{n} \frac{V_i}{\pi_i} (D_{si} - \theta_s) \quad s = 1, 2,$$

and for the parameter β_{jk} is

$$G_{\mathrm{IPW}}^{\beta_{jk}}(\alpha) = \sum_{i=1}^{n} g_{i,\mathrm{IPW}}^{\beta_{jk}}(\alpha) = \sum_{i=1}^{n} \frac{V_i}{\pi_i} (\mathrm{I}(T_i \geq c_j) D_{ki} - \beta_{jk}) \quad j = 1, 2;\ k = 1, 2, 3.$$

We show that these estimating functions are unbiased under assumptions (A1) and (A2). In fact, we get

$$\mathbb{E}\left[V_i \pi_i^{-1} (D_{si} - \theta_{s0}) \right] = \mathbb{E}_{D_s, T, A}\left[\mathbb{E}\left(V_i \pi_i^{-1} (D_{si} - \theta_{s0}) \,|, T_i, A_i \right) \right]$$

$$\begin{aligned} &= \mathbb{E}_{D_s,T,A}\left[\pi_i^{-1}\mathbb{E}\left(V_i|T_i,A_i\right)\left(\rho_{si}-\theta_{s0}\right)\right] \\ &= \mathbb{E}_{D_s,T,A}\left[\rho_{si}-\theta_{s0}\right]=0, \end{aligned}$$

and

$$\begin{aligned} &\mathbb{E}_{D_k,T_i,A_i}\left[\mathbb{E}\left[g_{i,\mathrm{IPW}}^{\beta_{jk}}(\alpha_0)|T_i,A_i\right]\right] \\ &= \mathbb{E}_{D_{ki},T_i,A_i}\left[\mathbb{E}\left[\left\{\frac{V_i}{\pi_i}\left(\mathrm{I}(T_i\geq c_j)D_{ki}-\beta_{jk0}\right)\right\}\middle|T_i,A_i\right]\right] \\ &= \mathbb{E}_{D_{ki},T_i,A_i}\left[\mathrm{I}(T_i\geq c_j)\rho_{ik}-\beta_{jk0}\right]=0. \end{aligned}$$

Therefore, condition (C1) holds for $G_{\mathrm{IPW}}^{\theta_s}(\alpha)$ and $G_{\mathrm{IPW}}^{\beta_{jk}}(\alpha)$, all s,j,k.

Next, we obtain the partial derivatives

$$\begin{aligned} \tfrac{\partial}{\partial\theta_{s'}}G_{\mathrm{IPW}}^{\theta_s}(\alpha) &= -\sum_{i=1}^{n}\frac{V_i}{\pi_i}\mathrm{I}(s'=s), & \tfrac{\partial}{\partial\beta_{jk}}G_{\mathrm{IPW}}^{\theta_s}(\alpha) &= 0, \\ \tfrac{\partial}{\partial\theta_s}G_{\mathrm{IPW}}^{\beta_{jk}}(\alpha) &= 0, & \tfrac{\partial}{\partial\beta_{j'k'}}G_{\mathrm{IPW}}^{\beta_{jk}}(\alpha) &= -\sum_{i=1}^{n}\frac{V_i}{\pi_i}\mathrm{I}(j'k'=jk), \end{aligned}$$

and, for the logistic model (used to estimate the verification process)

$$\begin{aligned} \frac{\partial}{\partial\tau_\pi}G_{\mathrm{IPW}}^{\theta_s}(\alpha) &= -\sum_{i=1}^{n}\frac{V_i(D_{si}-\theta_s)U_i}{e^{U_i^\top\tau_\pi}}, \\ \frac{\partial}{\partial\tau_\pi}G_{\mathrm{IPW}}^{\beta_{jk}}(\alpha) &= -\sum_{i=1}^{n}\frac{V_i(\mathrm{I}(T_i\geq c_j)D_{ki}-\beta_{jk})U_i}{e^{U_i^\top\tau_\pi}}, \end{aligned}$$

or the probit model

$$\begin{aligned} \frac{\partial}{\partial\tau_\pi}G_{\mathrm{IPW}}^{\theta_s}(\alpha) &= -\sum_{i=1}^{n}\frac{V_i(D_{si}-\theta_s)U_i\phi(U_i^\top\tau_\pi)}{\Phi^2(U_i^\top\tau_\pi)}, \\ \frac{\partial}{\partial\tau_\pi}G_{\mathrm{IPW}}^{\beta_{jk}}(\alpha) &= -\sum_{i=1}^{n}\frac{V_i(\mathrm{I}(T_i\geq c_j)D_{ki}-\beta_{jk})U_i\phi(U_i^\top\tau_\pi)}{\Phi^2(U_i^\top\tau_\pi)}. \end{aligned}$$

The computation of the second-order derivatives is similar and the results imply that the conditions (C2) and (C3) hold.

SPE estimator. Recall that

$$\begin{aligned} G_{\mathrm{SPE}}^{\theta_s}(\alpha) &= \sum_{i=1}^{n}\left\{\frac{V_i}{\pi_i}\left(D_{si}-\theta_s\right)-\frac{V_i-\pi_i}{\pi_i}\left(\rho_{si}-\theta_s\right)\right\}, \quad s=1,2, \\ G_{\mathrm{SPE}}^{\beta_{jk}}(\alpha) &= \sum_{i=1}^{n}\left\{\frac{V_i}{\pi_i}\left[\mathrm{I}(T_i\geq c_j)D_{ki}-\beta_{jk}\right]-\frac{V_i-\pi_i}{\pi_i}\left[\mathrm{I}(T_i\geq c_j)\rho_{ki}-\beta_{jk}\right]\right\}, \end{aligned}$$

for $j=1,2$ and $k=1,2,3$. Under assumption (A1), $\mathbb{E}\left[g_{i,\mathrm{SPE}}^{\theta_k}(\alpha_0)\right]$ equals

$$\mathbb{E}_{D_s,T_i,A_i}\left[\mathbb{E}\left[g_{i,\mathrm{SPE}}^{\theta_s}(\alpha_0)|T_i,A_i\right]\right]$$

$$
\begin{aligned}
&= \mathbb{E}_{D_{si},T_i,A_i}\left[\mathbb{E}\left[\left\{\frac{V_i}{\pi_i}\left(D_{si}-\theta_{s0}\right)-\frac{V_i-\pi_i}{\pi_i}\left(\rho_{si}-\theta_{s0}\right)\right\}|T_i,A_i\right]\right]\\
&= \mathbb{E}_{D_{si},T_i,A_i}\left[\pi_i^{-1}\left[\mathbb{E}\left[D_{si}|T_i,A_i\right]-\theta_{s0}\right]\pi_i\right]\\
&\quad -\pi_i^{-1}\mathbb{E}_{D_{si},T_i,A_i}\left[\mathbb{E}\left[(V_i-\pi_i)(\rho_{si}-\theta_{s0})|T_i,A_i\right]\right]\\
&= \mathbb{E}_{D_{si},T_i,A_i}\left[\rho_{si}-\theta_{s0}\right]-\pi_i^{-1}\mathbb{E}_{D_{si},T_i,A_i}\left[(\rho_{si}-\theta_{s0})\mathbb{E}\left[(V_i-\pi_i)|T_i,A_i\right]\right]\\
&= \mathbb{E}_{D_{si},T_i,A_i}\left[\rho_{si}-\theta_{s0}\right]=0.
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}_{D_k,T_i,A_i}\left[\mathbb{E}\left[g_{i,\mathrm{SPE}}^{\beta_{jk}}(\alpha_0)|T_i,A_i\right]\right]\\
&= \mathbb{E}_{D_{ki},T_i,A_i}\left[\mathbb{E}\left[\left\{\frac{V_i}{\pi_i}\left[\mathrm{I}(T_i\geq c_j)D_{ki}-\beta_{jk0}\right]\right.\right.\right.\\
&\quad \left.\left.\left.-\frac{V_i-\pi_i}{\pi_i}\left[\mathrm{I}(T_i\geq c_j)\rho_{ki}-\beta_{jk0}\right]\right\}\right|T_i,A_i\right]\right]\\
&= \mathbb{E}_{D_{ki},T_i,A_i}\left[\left[\mathrm{I}(T_i\geq c_j)\mathbb{E}\left[D_{ki}|T_i,A_i\right]-\beta_{jk0}\right]\right]\\
&\quad -\pi_i^{-1}\mathbb{E}_{D_{ki},T_i,A_i}\left[\mathbb{E}\left[(V_i-\pi_i)(\mathrm{I}(T_i\geq c_j)\rho_{ki}-\beta_{jk0})|T_i,A_i\right]\right]\\
&= \mathbb{E}_{D_{ki},T_i,A_i}\left[\mathrm{I}(T_i\geq c_j)\rho_{ki}-\beta_{jk0}\right]=0.
\end{aligned}
$$

Therefore, condition (C1) holds for $G_{\mathrm{SPE}}^{\theta_s}(\alpha)$ and $G_{\mathrm{SPE}}^{\beta_{jk}}(\alpha)$, all s,j,k.

Next, we obtain the partial derivatives

$$
\begin{array}{llll}
\frac{\partial}{\partial\theta_{s'}}G_{\mathrm{SPE}}^{\theta_s}(\alpha) & =-n\mathrm{I}(s'=s) & \frac{\partial}{\partial\beta_{jk}}G_{\mathrm{SPE}}^{\theta_s}(\alpha) & =0\\
\frac{\partial}{\partial\theta_s}G_{\mathrm{SPE}}^{\beta_{jk}}(\alpha) & =0 & \frac{\partial}{\partial\beta_{j'k'}}g_{\mathrm{SPE}}^{\beta_{jk}}(\alpha) & =-n\mathrm{I}(j'k'=jk)
\end{array}
$$

and the partial derivative with respect to $\tau_\rho\equiv(\tau_{\rho_1},\tau_{\rho_2})$

$$
\begin{aligned}
\frac{\partial}{\partial\tau_{\rho_l}}G_{\mathrm{SPE}}^{\theta_s}(\alpha) &= \sum_{i=1}^{n}-\frac{V_i-\pi_i}{\pi_i}\frac{\partial}{\partial\tau_{\rho_l}}\rho_{si},\\
\frac{\partial}{\partial\tau_{\rho_l}}G_{\mathrm{SPE}}^{\beta_{jk}}(\alpha) &= \sum_{i=1}^{n}-\frac{V_i-\pi_i}{\pi_i}\mathrm{I}(T_i\geq c_j)\frac{\partial}{\partial\tau_{\rho_l}}\rho_{si},
\end{aligned}
$$

where $\frac{\partial}{\partial\tau_{\rho_l}}\rho_{si}$ is given in (A.2). The partial derivative with respect to τ_π, are

$$
\begin{aligned}
\frac{\partial}{\partial\tau_\pi}G_{\mathrm{SPE}}^{\theta_s}(\alpha) &= \sum_{i=1}^{n}\frac{V_iU_i(\rho_{si}-D_{si})}{e^{U_i^\top\tau_\pi}},\\
\frac{\partial}{\partial\tau_\pi}G_{\mathrm{SPE}}^{\beta_{jk}}(\alpha) &= \sum_{i=1}^{n}\frac{V_iU_i\mathrm{I}(T_i\geq c_j)(\rho_{ki}-D_{ki})}{e^{U_i^\top\tau}},
\end{aligned}
$$

when the logistic model is used for the verification process. If the probit model is used, we have

$$
\frac{\partial}{\partial\tau_\pi}G_{\mathrm{SPE}}^{\theta_s}(\alpha) = \sum_{i=1}^{n}\frac{V_iU_i(D_{si}-\rho_{si})\phi(U_i^\top\tau_\pi)}{\Phi^2(U_i^\top\tau_\pi)},
$$

$$\frac{\partial}{\partial \tau_\pi} G^{\beta_{jk}}_{\mathrm{SPE}}(\alpha) \;=\; \sum_{i=1}^{n} \frac{V_i U_i \mathrm{I}(T_i \geq c_j)(D_{si} - \rho_{ki})\phi(U_i^\top \tau_\pi)}{\Phi^2(U_i^\top \tau_\pi)}.$$

Also in this case, computation of the second–order partial derivatives develops similarly and the results imply that the conditions (C2) and (C3) hold.

Asymptotic covariance matrix. Recall that the asymptotic covariance matrix of TCF estimators is obtained as

$$\frac{\partial h(\alpha_0)}{\partial \alpha^\top} \Sigma \frac{\partial h^\top(\alpha_0)}{\partial \alpha^\top},$$

where $h(\alpha) = \left(1 - \frac{\beta_{11}}{\theta_1}, \frac{\beta_{12}-\beta_{22}}{\theta_2}, \frac{\beta_{23}}{1-(\theta_1+\theta_2)}\right)^\top$ and

$$\Sigma = \left[\mathbb{E}\left\{\frac{\partial}{\partial \alpha^\top} g_{i,*}(\alpha_0)\right\}\right]^{-1} \mathbb{E}\{g_{i,*}(\alpha_0) g_{i,*}(\alpha_0)^\top\} \left[\mathbb{E}\left\{\frac{\partial}{\partial \alpha^\top} g_{i,*}^\top(\alpha_0)\right\}\right]^{-1}.$$

It is easy to derive that

$$\frac{\partial h(\alpha)}{\partial \alpha^\top} = \begin{pmatrix} \frac{\beta_{11}}{\theta_1^2} & 0 & -\frac{1}{\theta_1} & 0 & 0 & 0 & 0 \\ 0 & -\frac{\beta_{12}-\beta_{22}}{\theta_2^2} & 0 & \frac{1}{\theta_2} & -\frac{1}{\theta_2} & 0 & 0 \\ \frac{\beta_{23}}{(1-\theta_1-\theta_2)^2} & \frac{\beta_{23}}{(1-\theta_1-\theta_2)^2} & 0 & 0 & 0 & \frac{1}{1-\theta_1-\theta_2} & 0 \end{pmatrix}.$$

The elements $g_{i,*}(\alpha)$ of the estimating functions $G_*(\alpha)$ are given in the previous paragraphs. Now, we derive the explicit form for $\frac{\partial}{\partial \alpha^\top} g_{i,*}(\alpha)$.

First, we consider the class of imputation estimators. We get

$$\begin{aligned}
\frac{\partial}{\partial \alpha} g^{\theta_1}_{i,\mathrm{IE}}(\alpha) &= \left(-1,0,0,0,0,0,A_{11i}^\top, A_{21i}^\top\right)^\top, \\
\frac{\partial}{\partial \alpha} g^{\theta_2}_{i,\mathrm{IE}}(\alpha) &= \left(0,-1,0,0,0,0,A_{12i}^\top, A_{22i}^\top\right)^\top, \\
\frac{\partial}{\partial \alpha} g^{\beta_{11}}_{i,\mathrm{IE}}(\alpha) &= \left(0,0,-1,0,0,0,B_{111i}^\top, B_{121i}^\top\right)^\top, \\
\frac{\partial}{\partial \alpha} g^{\beta_{12}}_{i,\mathrm{IE}}(\alpha) &= \left(0,0,0,-1,0,0,B_{112i}^\top, B_{122i}^\top\right)^\top, \\
\frac{\partial}{\partial \alpha} g^{\beta_{22}}_{i,\mathrm{IE}}(\alpha) &= \left(0,0,0,0,-1,0,B_{212i}^\top, B_{222i}^\top\right)^\top, \\
\frac{\partial}{\partial \alpha} g^{\beta_{23}}_{i,\mathrm{IE}}(\alpha) &= \left(0,0,0,0,0,-1,B_{213i}^\top, B_{223i}^\top\right)^\top, \\
\frac{\partial}{\partial \alpha^\top} g^{\tau_{\rho_1}}_{i,\mathrm{IE}}(\alpha) &= \left(0,0,0,0,0,0,C_{11i}, C_{21i}\right), \\
\frac{\partial}{\partial \alpha^\top} g^{\tau_{\rho_2}}_{i,\mathrm{IE}}(\alpha) &= \left(0,0,0,0,0,0,C_{12i}, C_{22i}\right),
\end{aligned}$$

where

$$A_{lsi} = (1 - mV_i)\frac{\partial}{\partial \tau_{\rho_l}} \rho_{si}, \quad B_{jlki} = (1 - mV_i)\mathrm{I}(T_i \geq c_j)\frac{\partial}{\partial \tau_{\rho_l}} \rho_{ki},$$

$$C_{lsi} = \frac{\partial}{\partial \tau_{\rho_l}^\top} g_i^{\tau_{\rho_s}}(\alpha),$$

with $j, l, s = 1, 2$, $k = 1, 2, 3$ and $i = 1, \ldots, n$ (see (A.1) and (A.2) for the multinomial logistic model specification of the disease process).

Thus,

$$\frac{\partial}{\partial \alpha^\top} g_{i,\mathrm{IE}}(\alpha) = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 & A_{11i}^\top & A_{21i}^\top \\ 0 & -1 & 0 & 0 & 0 & 0 & A_{12i}^\top & A_{22i}^\top \\ 0 & 0 & -1 & 0 & 0 & 0 & B_{111i}^\top & B_{121i}^\top \\ 0 & 0 & 0 & -1 & 0 & 0 & B_{112i}^\top & B_{122i}^\top \\ 0 & 0 & 0 & 0 & -1 & 0 & B_{212i}^\top & B_{222i}^\top \\ 0 & 0 & 0 & 0 & 0 & -1 & B_{213i}^\top & B_{223i}^\top \\ 0 & 0 & 0 & 0 & 0 & 0 & C_{11i} & C_{21i} \\ 0 & 0 & 0 & 0 & 0 & 0 & C_{12i} & C_{22i} \end{pmatrix}.$$

Then, we consider the inverse probability weighted estimators. Let

$$A_{ki} = \frac{\partial}{\partial \tau_\pi} g_{i,\mathrm{IPW}}^{\theta_k}(\alpha), \qquad B_{jki} = \frac{\partial}{\partial \tau_\pi} g_{i,\mathrm{IPW}}^{\beta_{jk}}(\alpha), \qquad C_i = \frac{\partial}{\partial \tau_\pi^\top} g_i^{\tau_\pi}(\alpha).$$

Note that these quantities change according to the model, logit or probit, chosen for the verification process. We obtain

$$\begin{aligned}
\frac{\partial}{\partial \alpha} g_{i,\mathrm{IPW}}^{\theta_1}(\alpha) &= \left(-V_i \pi_i^{-1}, 0, 0, 0, 0, 0, A_{1i}^\top\right)^\top, \\
\frac{\partial}{\partial \alpha} g_{i,\mathrm{IPW}}^{\theta_2}(\alpha) &= \left(0, -V_i \pi_i^{-1}, 0, 0, 0, 0, A_{2i}^\top\right)^\top, \\
\frac{\partial}{\partial \alpha} g_{i,\mathrm{IPW}}^{\beta_{11}}(\alpha) &= \left(0, 0, -V_i \pi_i^{-1}, 0, 0, 0, B_{11i}^\top\right)^\top, \\
\frac{\partial}{\partial \alpha} g_{i,\mathrm{IPW}}^{\beta_{12}}(\alpha) &= \left(0, 0, 0, -V_i \pi_i^{-1}, 0, 0, B_{12i}^\top\right)^\top, \\
\frac{\partial}{\partial \alpha} g_{i,\mathrm{IPW}}^{\beta_{22}}(\alpha) &= \left(0, 0, 0, 0, -V_i \pi_i^{-1}, 0, B_{22i}^\top\right)^\top, \\
\frac{\partial}{\partial \alpha} g_{i,\mathrm{IPW}}^{\beta_{23}}(\alpha) &= \left(0, 0, 0, 0, 0, -V_i \pi_i^{-1}, B_{23i}^\top\right)^\top, \\
\frac{\partial}{\partial \alpha^\top} g_i^{\tau_\pi}(\alpha) &= \left(0, 0, 0, 0, 0, 0, C_i\right),
\end{aligned}$$

Summarizing

$$\frac{\partial}{\partial \alpha^\top} g_{i,\mathrm{IPW}}(\alpha) = \begin{pmatrix} -\frac{V_i}{\pi_i} & 0 & 0 & 0 & 0 & 0 & A_{1i}^\top \\ 0 & -\frac{V_i}{\pi_i} & 0 & 0 & 0 & 0 & A_{2i}^\top \\ 0 & 0 & -\frac{V_i}{\pi_i} & 0 & 0 & 0 & B_{11i}^\top \\ 0 & 0 & 0 & -\frac{V_i}{\pi_i} & 0 & 0 & B_{12i}^\top \\ 0 & 0 & 0 & 0 & -\frac{V_i}{\pi_i} & 0 & B_{22i}^\top \\ 0 & 0 & 0 & 0 & 0 & -\frac{V_i}{\pi_i} & B_{23i}^\top \\ 0 & 0 & 0 & 0 & 0 & 0 & C_i \end{pmatrix}.$$

Finally, we consider the SPE estimators. We have

$$\begin{aligned}
\frac{\partial}{\partial\alpha} g_{i,\mathrm{SPE}}^{\theta_1}(\alpha) &= \left(-1,0,0,0,0,0,H_{11i}^\top,H_{21i}^\top,D_{1i}^\top\right)^\top,\\
\frac{\partial}{\partial\alpha} g_{i,\mathrm{SPE}}^{\theta_2}(\alpha) &= \left(0,-1,0,0,0,0,H_{12i}^\top,H_{22i}^\top,D_{2i}^\top\right)^\top,\\
\frac{\partial}{\partial\alpha} g_{i,\mathrm{SPE}}^{\beta_{11}}(\alpha) &= \left(0,0,-1,0,0,0,G_{111i}^\top,G_{121i}^\top,E_{11i}^\top\right)^\top,\\
\frac{\partial}{\partial\alpha} g_{i,\mathrm{SPE}}^{\beta_{12}}(\alpha) &= \left(0,0,0,-1,0,0,G_{112i}^\top,G_{122i}^\top,E_{12i}^\top\right)^\top,\\
\frac{\partial}{\partial\alpha} g_{i,\mathrm{SPE}}^{\beta_{22}}(\alpha) &= \left(0,0,0,0,-1,0,G_{212i}^\top,G_{222i}^\top,E_{22i}^\top\right)^\top,\\
\frac{\partial}{\partial\alpha} g_{i,\mathrm{SPE}}^{\beta_{23}}(\alpha) &= \left(0,0,0,0,0,-1,G_{213i}^\top,G_{223i}^\top,E_{23i}^\top\right)^\top,\\
\frac{\partial}{\partial\alpha^\top} g_{i,\mathrm{SPE}}^{\tau_{\rho_1}}(\alpha) &= \left(0,0,0,0,0,0,C_{11i},C_{21i},0\right),\\
\frac{\partial}{\partial\alpha^\top} g_{i,\mathrm{SPE}}^{\tau_{\rho_2}}(\alpha) &= \left(0,0,0,0,0,0,C_{12i},C_{22i},0\right),\\
\frac{\partial}{\partial\alpha^\top} g_{i,\mathrm{SPE}}^{\tau_{\pi}}(\alpha) &= \left(0,0,0,0,0,0,0,0,C_i\right),
\end{aligned}$$

where

$$\begin{aligned}
H_{lki} &= -\frac{V_i-\pi_i}{\pi_i}\frac{\partial}{\partial\tau_{\rho_l}}\rho_{ki}, \quad & G_{jlki} &= -\tfrac{V_i-\pi_i}{\pi_i}\mathrm{I}(T_i\ge c_j)\tfrac{\partial}{\partial\tau_{\rho_l}}\rho_{ki},\\
D_{si} &= \frac{\partial}{\partial\tau_\pi} g_{i,\mathrm{SPE}}^{\theta_s}(\alpha), & E_{jki} &= \frac{\partial}{\partial\tau_\pi} g_{i,\mathrm{SPE}}^{\beta_{jk}}(\alpha),
\end{aligned}$$

and C_{lsi} and C_i are defined above. Therefore

$$\frac{\partial}{\partial\alpha^\top} g_{i,\mathrm{SPE}}(\alpha) = \begin{pmatrix}
-1 & 0 & 0 & 0 & 0 & 0 & H_{11i}^\top & H_{21i}^\top & D_{1i}^\top\\
0 & -1 & 0 & 0 & 0 & 0 & H_{12i}^\top & H_{22i}^\top & D_{2i}^\top\\
0 & 0 & -1 & 0 & 0 & 0 & G_{111i}^\top & G_{121i}^\top & E_{11i}^\top\\
0 & 0 & 0 & -1 & 0 & 0 & G_{112i}^\top & G_{122i}^\top & E_{12i}^\top\\
0 & 0 & 0 & 0 & -1 & 0 & G_{212i}^\top & G_{222i}^\top & E_{22i}^\top\\
0 & 0 & 0 & 0 & 0 & -1 & G_{213i}^\top & G_{223i}^\top & E_{23i}^\top\\
0 & 0 & 0 & 0 & 0 & 0 & C_{11i} & C_{21i} & 0\\
0 & 0 & 0 & 0 & 0 & 0 & C_{12i} & C_{22i} & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & C_i
\end{pmatrix}.$$

Appendix B: Simulation results of Study 1 and Study 2

In this section, we present results of simulations in Study 1 and Study 2. Tables 9–14 show simulation results of Study 1 for sample sizes equal to 500 and 1000, respectively. The results of Study 2 are presented in Tables 15 and 16, corresponding to the first and third value of Λ, respectively.

TABLE 9. *Simulation results from 5000 replications when both models for ρ_k and π are correctly specified (Study 1) and the first value of Λ is considered. "True" denotes the true parameter value. Sample size = 500.*

	TCF_1	TCF_2	TCF_3	$MC.sd_1$	$MC.sd_2$	$MC.sd_3$	$asy.sd_1$	$asy.sd_2$	$asy.sd_3$	$boot.sd_1$	$boot.sd_2$	$boot.sd_3$
						cut-point = (2,4)						
True	0.5000	0.4347	0.9347									
FI	0.5007	0.4357	0.9343	0.0373	0.0339	0.0190	0.0343	0.0309	0.0360	0.0372	0.0340	0.0193
MSI	0.5008	0.4357	0.9343	0.0382	0.0378	0.0225	0.0353	0.0354	0.0380	0.0381	0.0380	0.0227
IPW	0.5020	0.4357	0.9345	0.0506	0.0508	0.0260	0.0493	0.0502	0.0280	0.0499	0.0507	0.0262
SPE	0.5012	0.4357	0.9345	0.0401	0.0456	0.0256	0.0399	0.0450	0.0249	0.0409	0.0457	0.0259
						cut-point = (2,5)						
True	0.5000	0.7099	0.7752									
FI	0.5007	0.7115	0.7747	0.0373	0.0329	0.0372	0.0343	0.0321	0.0436	0.0372	0.0329	0.0374
MSI	0.5008	0.7111	0.7743	0.0382	0.0361	0.0395	0.0353	0.0355	0.0455	0.0381	0.0362	0.0396
IPW	0.5020	0.7111	0.7743	0.0506	0.0496	0.0464	0.0493	0.0479	0.0500	0.0499	0.0487	0.0463
SPE	0.5012	0.7112	0.7744	0.0401	0.0434	0.0442	0.0399	0.0425	0.0433	0.0409	0.0436	0.0448
						cut-point = (2,7)						
True	0.5000	0.9230	0.2248									
FI	0.5007	0.9228	0.2241	0.0373	0.0167	0.0377	0.0343	0.0229	0.0310	0.0372	0.0169	0.0370
MSI	0.5008	0.9230	0.2242	0.0382	0.0199	0.0382	0.0353	0.0253	0.0317	0.0381	0.0202	0.0376
IPW	0.5020	0.9232	0.2242	0.0506	0.0266	0.0534	0.0493	0.0251	0.0520	0.0499	0.0263	0.0525
SPE	0.5012	0.9235	0.2245	0.0401	0.0255	0.0416	0.0399	0.0244	0.0403	0.0409	0.0253	0.0545
						cut-point = (4,5)						
True	0.9347	0.2752	0.7752									
FI	0.9349	0.2758	0.7747	0.0176	0.0285	0.0372	0.0161	0.0246	0.0436	0.0174	0.0289	0.0374
MSI	0.9348	0.2754	0.7743	0.0194	0.0326	0.0395	0.0179	0.0291	0.0455	0.0191	0.0328	0.0396
IPW	0.9352	0.2754	0.7743	0.0299	0.0472	0.0464	0.0263	0.0466	0.0500	0.0284	0.0472	0.0463
SPE	0.9353	0.2755	0.7744	0.0270	0.0407	0.0442	0.0249	0.0399	0.0433	0.0291	0.0404	0.0448
						cut-point = (4,7)						
True	0.9347	0.4883	0.2248									
FI	0.9349	0.4872	0.2241	0.0176	0.0375	0.0377	0.0161	0.0347	0.0310	0.0174	0.0375	0.0370
MSI	0.9348	0.4872	0.2242	0.0194	0.0396	0.0382	0.0179	0.0373	0.0317	0.0191	0.0398	0.0376
IPW	0.9352	0.4876	0.2242	0.0299	0.0520	0.0534	0.0263	0.0511	0.0520	0.0284	0.0516	0.0525
SPE	0.9353	0.4877	0.2245	0.0270	0.0462	0.0416	0.0249	0.0456	0.0403	0.0291	0.0463	0.0545
						cut-point = (5,7)						
True	0.9883	0.2132	0.2248									
FI	0.9882	0.2114	0.2241	0.0051	0.0310	0.0377	0.0047	0.0274	0.0310	0.0054	0.0306	0.0370
MSI	0.9882	0.2118	0.2242	0.0069	0.0330	0.0382	0.0058	0.0299	0.0317	0.0069	0.0329	0.0376
IPW	0.9886	0.2121	0.2242	0.0137	0.0467	0.0534	0.0088	0.0441	0.0520	0.0130	0.0452	0.0525
SPE	0.9886	0.2123	0.2245	0.0133	0.0398	0.0416	0.0097	0.0387	0.0403	0.0127	0.0398	0.0545

TABLE 10. *Simulation results from 5000 replications when both models for ρ_k and π are correctly specified (Study 1) and the first value of Λ is considered. "True" denotes the true parameter value. Sample size = 1000.*

	TCF_1	TCF_2	TCF_3	$MC.sd_1$	$MC.sd_2$	$MC.sd_3$	$asy.sd_1$	$asy.sd_2$	$asy.sd_3$	$boot.sd_1$	$boot.sd_2$	$boot.sd_3$
						cut-point = (2,4)						
True	0.5000	0.4347	0.9347									
FI	0.5001	0.4346	0.9348	0.0265	0.0235	0.0133	0.0242	0.0217	0.0254	0.0262	0.0238	0.0135
MSI	0.5002	0.4349	0.9349	0.0273	0.0264	0.0157	0.0250	0.0250	0.0268	0.0269	0.0268	0.0160
IPW	0.5006	0.4357	0.9349	0.0362	0.0357	0.0184	0.0352	0.0358	0.0198	0.0353	0.0360	0.0185
SPE	0.5004	0.4353	0.9349	0.0287	0.0321	0.0180	0.0282	0.0320	0.0180	0.0283	0.0322	0.0182
						cut-point = (2,5)						
True	0.5000	0.7099	0.7752									
FI	0.5001	0.7096	0.7758	0.0265	0.0232	0.0260	0.0242	0.0227	0.0308	0.0262	0.0232	0.0263
MSI	0.5002	0.7095	0.7756	0.0273	0.0256	0.0276	0.0250	0.0251	0.0321	0.0269	0.0256	0.0279
IPW	0.5006	0.7104	0.7756	0.0362	0.0349	0.0325	0.0352	0.0342	0.0354	0.0353	0.0345	0.0327
SPE	0.5004	0.7100	0.7757	0.0287	0.0309	0.0307	0.0282	0.0303	0.0308	0.0283	0.0305	0.0310
						cut-point = (2,7)						
True	0.5000	0.9230	0.2248									
FI	0.5001	0.9228	0.2250	0.0265	0.0117	0.0260	0.0242	0.0160	0.0220	0.0262	0.0119	0.0262
MSI	0.5002	0.9230	0.2252	0.0273	0.0141	0.0265	0.0250	0.0178	0.0226	0.0269	0.0142	0.0266
IPW	0.5006	0.9233	0.2258	0.0362	0.0187	0.0383	0.0352	0.0181	0.0374	0.0353	0.0186	0.0375
SPE	0.5004	0.9235	0.2256	0.0287	0.0180	0.0286	0.0282	0.0176	0.0286	0.0283	0.0180	0.0291
						cut-point = (4,5)						
True	0.9347	0.2752	0.7752									
FI	0.9346	0.2749	0.7758	0.0124	0.0203	0.0260	0.0115	0.0173	0.0308	0.0123	0.0203	0.0263
MSI	0.9345	0.2746	0.7756	0.0137	0.0232	0.0276	0.0128	0.0205	0.0321	0.0136	0.0231	0.0279
IPW	0.9346	0.2748	0.7756	0.0213	0.0337	0.0325	0.0196	0.0332	0.0354	0.0205	0.0335	0.0327
SPE	0.9344	0.2747	0.7757	0.0190	0.0286	0.0307	0.0183	0.0283	0.0308	0.0187	0.0285	0.0310
						cut-point = (4,7)						
True	0.9347	0.4883	0.2248									
FI	0.9346	0.4882	0.2250	0.0124	0.0262	0.0260	0.0115	0.0245	0.0220	0.0123	0.0264	0.0262
MSI	0.9345	0.4881	0.2252	0.0137	0.0279	0.0265	0.0128	0.0263	0.0226	0.0136	0.0280	0.0266
IPW	0.9346	0.4876	0.2258	0.0213	0.0365	0.0383	0.0196	0.0364	0.0374	0.0205	0.0366	0.0375
SPE	0.9344	0.4882	0.2256	0.0190	0.0325	0.0286	0.0183	0.0324	0.0286	0.0187	0.0326	0.0291
						cut-point = (5,7)						
True	0.9883	0.2132	0.2248									
FI	0.9881	0.2132	0.2250	0.0036	0.0217	0.0260	0.0033	0.0194	0.0220	0.0037	0.0216	0.0262
MSI	0.9881	0.2135	0.2252	0.0048	0.0234	0.0265	0.0044	0.0212	0.0226	0.0049	0.0232	0.0266
IPW	0.9882	0.2129	0.2258	0.0100	0.0325	0.0383	0.0077	0.0317	0.0374	0.0097	0.0320	0.0375
SPE	0.9880	0.2135	0.2256	0.0097	0.0282	0.0286	0.0080	0.0276	0.0286	0.0094	0.0278	0.0291

TABLE 11. *Simulation results from 5000 replications when both models for ρ_k and π are correctly specified (Study 1) and the second value of Λ is considered. "True" denotes the true parameter value. Sample size = 500.*

	TCF_1	TCF_2	TCF_3	$MC.sd_1$	$MC.sd_2$	$MC.sd_3$	$asy.sd_1$	$asy.sd_2$	$asy.sd_3$	$boot.sd_1$	$boot.sd_2$	$boot.sd_3$
						cut-point = (2,4)						
True	0.5000	0.3970	0.8970									
FI	0.4999	0.3974	0.8965	0.0356	0.0294	0.0253	0.0326	0.0263	0.0355	0.0355	0.0298	0.0256
MSI	0.4999	0.3975	0.8961	0.0368	0.0355	0.0291	0.0339	0.0326	0.0380	0.0367	0.0355	0.0291
IPW	0.5000	0.3977	0.8962	0.0470	0.0492	0.0373	0.0460	0.0484	0.0369	0.0464	0.0487	0.0368
SPE	0.5000	0.3976	0.8963	0.0402	0.0446	0.0363	0.0397	0.0438	0.0348	0.0400	0.0442	0.0356
						cut-point = (2,5)						
True	0.5000	0.6335	0.7365									
FI	0.4999	0.6342	0.7360	0.0356	0.0303	0.0410	0.0326	0.0287	0.0439	0.0355	0.0306	0.0409
MSI	0.4999	0.6339	0.7358	0.0368	0.0356	0.0437	0.0339	0.0342	0.0463	0.0367	0.0357	0.0435
IPW	0.5000	0.6336	0.7363	0.0470	0.0477	0.0528	0.0460	0.0471	0.0529	0.0464	0.0474	0.0514
SPE	0.5000	0.6341	0.7362	0.0402	0.0440	0.0494	0.0397	0.0434	0.0479	0.0400	0.0437	0.0483
						cut-point = (2,7)						
True	0.5000	0.8682	0.2635									
FI	0.4999	0.8677	0.2631	0.0356	0.0222	0.0388	0.0326	0.0233	0.0352	0.0355	0.0219	0.0391
MSI	0.4999	0.8678	0.2633	0.0368	0.0263	0.0401	0.0339	0.0272	0.0370	0.0367	0.0261	0.0407
IPW	0.5000	0.8677	0.2638	0.0470	0.0354	0.0477	0.0460	0.0341	0.0486	0.0464	0.0349	0.0484
SPE	0.5000	0.8679	0.2635	0.0402	0.0336	0.0420	0.0397	0.0326	0.0420	0.0400	0.0331	0.0424
						cut-point = (4,5)						
True	0.8970	0.2365	0.7365									
FI	0.8972	0.2368	0.7360	0.0205	0.0257	0.0410	0.0195	0.0223	0.0439	0.0203	0.0258	0.0409
MSI	0.8968	0.2364	0.7358	0.0229	0.0310	0.0437	0.0219	0.0279	0.0463	0.0226	0.0308	0.0435
IPW	0.8969	0.2359	0.7363	0.0268	0.0421	0.0528	0.0261	0.0411	0.0529	0.0265	0.0415	0.0514
SPE	0.8967	0.2365	0.7362	0.0260	0.0374	0.0494	0.0254	0.0370	0.0479	0.0257	0.0373	0.0483
						cut-point = (4,7)						
True	0.8970	0.4711	0.2635									
FI	0.8972	0.4703	0.2631	0.0205	0.0356	0.0388	0.0195	0.0328	0.0352	0.0203	0.0356	0.0391
MSI	0.8968	0.4703	0.2633	0.0229	0.0398	0.0401	0.0219	0.0370	0.0370	0.0226	0.0394	0.0407
IPW	0.8969	0.4699	0.2638	0.0268	0.0492	0.0477	0.0261	0.0483	0.0486	0.0265	0.0486	0.0484
SPE	0.8967	0.4703	0.2635	0.0260	0.0454	0.0420	0.0254	0.0445	0.0420	0.0257	0.0449	0.0424
						cut-point = (5,7)						
True	0.9711	0.2347	0.2635									
FI	0.9710	0.2335	0.2631	0.0086	0.0283	0.0388	0.0084	0.0260	0.0352	0.0088	0.0284	0.0391
MSI	0.9711	0.2339	0.2633	0.0116	0.0327	0.0401	0.0111	0.0304	0.0370	0.0117	0.0325	0.0407
IPW	0.9711	0.2341	0.2638	0.0144	0.0402	0.0477	0.0136	0.0397	0.0486	0.0143	0.0400	0.0484
SPE	0.9711	0.2339	0.2635	0.0142	0.0376	0.0420	0.0135	0.0370	0.0420	0.0141	0.0373	0.0424

TABLE 12. *Simulation results from 5000 replications when both models for ρ_k and π are correctly specified (Study 1) and the second value of Λ is considered. "True" denotes the true parameter value. Sample size = 1000.*

	TCF_1	TCF_2	TCF_3	$MC.sd_1$	$MC.sd_2$	$MC.sd_3$	$asy.sd_1$	$asy.sd_2$	$asy.sd_3$	$boot.sd_1$	$boot.sd_2$	$boot.sd_3$
						cut-point = (2,4)						
True	0.5000	0.3970	0.8970									
FI	0.4997	0.3967	0.8966	0.0248	0.0208	0.0177	0.0230	0.0185	0.0250	0.0250	0.0209	0.0180
MSI	0.4997	0.3965	0.8966	0.0257	0.0251	0.0202	0.0240	0.0230	0.0268	0.0259	0.0250	0.0205
IPW	0.4994	0.3967	0.8967	0.0323	0.0349	0.0259	0.0327	0.0343	0.0263	0.0327	0.0344	0.0260
SPE	0.4997	0.3966	0.8967	0.0279	0.0317	0.0251	0.0281	0.0311	0.0250	0.0282	0.0311	0.0252
						cut-point = (2,5)						
True	0.5000	0.6335	0.7365									
FI	0.4997	0.6330	0.7364	0.0248	0.0215	0.0286	0.0230	0.0203	0.0310	0.0250	0.0216	0.0288
MSI	0.4997	0.6327	0.7361	0.0257	0.0253	0.0304	0.0240	0.0241	0.0327	0.0259	0.0252	0.0307
IPW	0.4994	0.6326	0.7365	0.0323	0.0339	0.0360	0.0327	0.0335	0.0375	0.0327	0.0335	0.0363
SPE	0.4997	0.6328	0.7362	0.0279	0.0314	0.0338	0.0281	0.0308	0.0340	0.0282	0.0309	0.0341
						cut-point = (2,7)						
True	0.5000	0.8682	0.2635									
FI	0.4997	0.8679	0.2640	0.0248	0.0153	0.0274	0.0230	0.0164	0.0249	0.0250	0.0154	0.0275
MSI	0.4997	0.8680	0.2643	0.0257	0.0183	0.0286	0.0240	0.0192	0.0262	0.0259	0.0184	0.0287
IPW	0.4994	0.8682	0.2645	0.0323	0.0248	0.0343	0.0327	0.0244	0.0345	0.0327	0.0246	0.0341
SPE	0.4997	0.8682	0.2644	0.0279	0.0236	0.0299	0.0281	0.0232	0.0297	0.0282	0.0234	0.0298
						cut-point = (4,5)						
True	0.8970	0.2365	0.7365									
FI	0.8971	0.2363	0.7364	0.0144	0.0180	0.0286	0.0138	0.0157	0.0310	0.0143	0.0182	0.0288
MSI	0.8971	0.2362	0.7361	0.0160	0.0217	0.0304	0.0155	0.0197	0.0327	0.0160	0.0217	0.0307
IPW	0.8972	0.2359	0.7365	0.0188	0.0297	0.0360	0.0186	0.0291	0.0375	0.0187	0.0293	0.0363
SPE	0.8972	0.2362	0.7362	0.0183	0.0264	0.0338	0.0181	0.0262	0.0340	0.0182	0.0262	0.0341
						cut-point = (4,7)						
True	0.8970	0.4711	0.2635									
FI	0.8971	0.4712	0.2640	0.0144	0.0252	0.0274	0.0138	0.0232	0.0249	0.0143	0.0250	0.0275
MSI	0.8971	0.4715	0.2643	0.0160	0.0280	0.0286	0.0155	0.0261	0.0262	0.0160	0.0278	0.0287
IPW	0.8972	0.4715	0.2645	0.0188	0.0348	0.0343	0.0186	0.0342	0.0345	0.0187	0.0343	0.0341
SPE	0.8972	0.4717	0.2644	0.0183	0.0321	0.0299	0.0181	0.0316	0.0297	0.0182	0.0316	0.0298
						cut-point = (5,7)						
True	0.9711	0.2347	0.2635									
FI	0.9709	0.2350	0.2640	0.0061	0.0201	0.0274	0.0060	0.0184	0.0249	0.0062	0.0200	0.0275
MSI	0.9709	0.2353	0.2643	0.0082	0.0229	0.0286	0.0080	0.0216	0.0262	0.0082	0.0229	0.0287
IPW	0.9709	0.2356	0.2645	0.0101	0.0285	0.0343	0.0099	0.0282	0.0345	0.0102	0.0283	0.0341
SPE	0.9710	0.2354	0.2644	0.0100	0.0266	0.0299	0.0098	0.0263	0.0297	0.0100	0.0264	0.0298

TABLE 13. *Simulation results from 5000 replications when both models for ρ_k and π are correctly specified (Study 1) and the third value of Λ is considered. "True" denotes the true parameter value. Sample size = 500.*

	TCF_1	TCF_2	TCF_3	$MC.sd_1$	$MC.sd_2$	$MC.sd_3$	$asy.sd_1$	$asy.sd_2$	$asy.sd_3$	$boot.sd_1$	$boot.sd_2$	$boot.sd_3$
						cut-point = (2,4)						
True	0.5000	0.3031	0.8031									
FI	0.5001	0.3027	0.8034	0.0356	0.0240	0.0348	0.0320	0.0206	0.0373	0.0350	0.0242	0.0346
MSI	0.5001	0.3031	0.8034	0.0375	0.0322	0.0384	0.0340	0.0291	0.0408	0.0367	0.0318	0.0383
IPW	0.5004	0.3031	0.8037	0.0454	0.0444	0.0454	0.0438	0.0439	0.0451	0.0440	0.0441	0.0452
SPE	0.5002	0.3032	0.8036	0.0410	0.0411	0.0441	0.0400	0.0406	0.0438	0.0401	0.0407	0.0440
						cut-point = (2,5)						
True	0.5000	0.4682	0.6651									
FI	0.5001	0.4681	0.6656	0.0356	0.0266	0.0438	0.0320	0.0237	0.0430	0.0350	0.0271	0.0428
MSI	0.5001	0.4679	0.6654	0.0375	0.0348	0.0469	0.0340	0.0325	0.0461	0.0367	0.0350	0.0459
IPW	0.5004	0.4676	0.6655	0.0454	0.0475	0.0538	0.0438	0.0474	0.0526	0.0440	0.0476	0.0524
SPE	0.5002	0.4678	0.6654	0.0410	0.0440	0.0513	0.0400	0.0440	0.0500	0.0401	0.0442	0.0503
						cut-point = (2,7)						
True	0.5000	0.7027	0.3349									
FI	0.5001	0.7033	0.3346	0.0356	0.0268	0.0424	0.0320	0.0246	0.0383	0.0350	0.0267	0.0412
MSI	0.5001	0.7033	0.3346	0.0375	0.0336	0.0455	0.0340	0.0318	0.0414	0.0367	0.0334	0.0441
IPW	0.5004	0.7031	0.3352	0.0454	0.0439	0.0515	0.0438	0.0437	0.0505	0.0440	0.0440	0.0504
SPE	0.5002	0.7034	0.3347	0.0410	0.0416	0.0481	0.0400	0.0413	0.0465	0.0401	0.0414	0.0468
						cut-point = (4,5)						
True	0.8031	0.1651	0.6651									
FI	0.8033	0.1654	0.6656	0.0278	0.0196	0.0438	0.0260	0.0166	0.0430	0.0274	0.0196	0.0428
MSI	0.8030	0.1648	0.6654	0.0303	0.0256	0.0469	0.0284	0.0236	0.0461	0.0297	0.0259	0.0459
IPW	0.8030	0.1645	0.6655	0.0344	0.0346	0.0538	0.0335	0.0346	0.0526	0.0337	0.0349	0.0524
SPE	0.8030	0.1645	0.6654	0.0334	0.0317	0.0513	0.0325	0.0321	0.0500	0.0326	0.0322	0.0503
						cut-point = (4,7)						
True	0.8031	0.3996	0.3349									
FI	0.8033	0.4007	0.3346	0.0278	0.0300	0.0424	0.0260	0.0268	0.0383	0.0274	0.0299	0.0412
MSI	0.8030	0.4002	0.3346	0.0303	0.0367	0.0455	0.0284	0.0339	0.0414	0.0297	0.0364	0.0441
IPW	0.8030	0.4000	0.3352	0.0344	0.0458	0.0515	0.0335	0.0456	0.0505	0.0337	0.0458	0.0504
SPE	0.8030	0.4002	0.3347	0.0334	0.0431	0.0481	0.0325	0.0429	0.0465	0.0326	0.0430	0.0468
						cut-point = (5,7)						
True	0.8996	0.2345	0.3349									
FI	0.8996	0.2353	0.3346	0.0192	0.0245	0.0424	0.0182	0.0220	0.0383	0.0190	0.0248	0.0412
MSI	0.8996	0.2354	0.3346	0.0221	0.0307	0.0455	0.0212	0.0288	0.0414	0.0219	0.0310	0.0441
IPW	0.8997	0.2355	0.3352	0.0253	0.0384	0.0515	0.0249	0.0388	0.0505	0.0252	0.0391	0.0504
SPE	0.8997	0.2356	0.3347	0.0249	0.0364	0.0481	0.0246	0.0366	0.0465	0.0247	0.0367	0.0468

TABLE 14. *Simulation results from 5000 replications when both models for ρ_k and π are correctly specified (Study 1) and the third value of Λ is considered. "True" denotes the true parameter value. Sample size = 1000.*

	TCF_1	TCF_2	TCF_3	$MC.sd_1$	$MC.sd_2$	$MC.sd_3$	$asy.sd_1$	$asy.sd_2$	$asy.sd_3$	$boot.sd_1$	$boot.sd_2$	$boot.sd_3$
						cut-point = (2,4)						
True	0.5000	0.3031	0.8031									
FI	0.5003	0.3030	0.8040	0.0242	0.0169	0.0243	0.0226	0.0145	0.0264	0.0247	0.0170	0.0243
MSI	0.5001	0.3030	0.8038	0.0256	0.0222	0.0270	0.0240	0.0206	0.0288	0.0259	0.0224	0.0270
IPW	0.5001	0.3032	0.8038	0.0310	0.0310	0.0320	0.0310	0.0311	0.0320	0.0311	0.0311	0.0319
SPE	0.5001	0.3030	0.8040	0.0281	0.0285	0.0312	0.0283	0.0288	0.0310	0.0283	0.0287	0.0311
						cut-point = (2,5)						
True	0.5000	0.4682	0.6651									
FI	0.5003	0.4682	0.6663	0.0242	0.0193	0.0301	0.0226	0.0167	0.0304	0.0247	0.0191	0.0301
MSI	0.5001	0.4681	0.6663	0.0256	0.0248	0.0320	0.0240	0.0230	0.0326	0.0259	0.0247	0.0324
IPW	0.5001	0.4683	0.6664	0.0310	0.0337	0.0368	0.0310	0.0336	0.0373	0.0311	0.0336	0.0370
SPE	0.5001	0.4682	0.6665	0.0281	0.0311	0.0350	0.0283	0.0312	0.0355	0.0283	0.0312	0.0355
						cut-point = (2,7)						
True	0.5000	0.7027	0.3349									
FI	0.5003	0.7028	0.3359	0.0242	0.0188	0.0289	0.0226	0.0173	0.0271	0.0247	0.0188	0.0290
MSI	0.5001	0.7025	0.3359	0.0256	0.0236	0.0307	0.0240	0.0225	0.0293	0.0259	0.0236	0.0311
IPW	0.5001	0.7023	0.3360	0.0310	0.0311	0.0350	0.0310	0.0310	0.0358	0.0311	0.0311	0.0356
SPE	0.5001	0.7024	0.3358	0.0281	0.0292	0.0324	0.0283	0.0293	0.0329	0.0283	0.0293	0.0330
						cut-point = (4,5)						
True	0.8031	0.1651	0.6651									
FI	0.8034	0.1652	0.6663	0.0193	0.0139	0.0301	0.0184	0.0117	0.0304	0.0193	0.0138	0.0301
MSI	0.8032	0.1652	0.6663	0.0211	0.0184	0.0320	0.0201	0.0167	0.0326	0.0209	0.0183	0.0324
IPW	0.8034	0.1651	0.6664	0.0241	0.0248	0.0368	0.0237	0.0246	0.0373	0.0237	0.0247	0.0370
SPE	0.8032	0.1653	0.6665	0.0233	0.0229	0.0350	0.0230	0.0228	0.0355	0.0230	0.0228	0.0355
						cut-point = (4,7)						
True	0.8031	0.3996	0.3349									
FI	0.8034	0.3998	0.3359	0.0193	0.0207	0.0289	0.0184	0.0189	0.0271	0.0193	0.0210	0.0290
MSI	0.8032	0.3995	0.3359	0.0211	0.0253	0.0307	0.0201	0.0240	0.0293	0.0209	0.0256	0.0311
IPW	0.8034	0.3991	0.3360	0.0241	0.0319	0.0350	0.0237	0.0323	0.0358	0.0237	0.0323	0.0356
SPE	0.8032	0.3994	0.3358	0.0233	0.0299	0.0324	0.0230	0.0303	0.0329	0.0230	0.0304	0.0330
						cut-point = (5,7)						
True	0.8996	0.2345	0.3349									
FI	0.8998	0.2346	0.3359	0.0134	0.0172	0.0289	0.0129	0.0155	0.0271	0.0134	0.0174	0.0290
MSI	0.8997	0.2343	0.3359	0.0157	0.0216	0.0307	0.0150	0.0204	0.0293	0.0155	0.0218	0.0311
IPW	0.8998	0.2340	0.3360	0.0180	0.0273	0.0350	0.0177	0.0274	0.0358	0.0177	0.0275	0.0356
SPE	0.8997	0.2342	0.3358	0.0178	0.0256	0.0324	0.0174	0.0258	0.0329	0.0174	0.0259	0.0330

TABLE 15. *Simulation results from 5000 replications when the model for the verification process is misspecified (Study 2) and the first value of* Λ *is used. "True" indicates the true parameter value. Sample size = 1000.*

	TCF_1	TCF_2	TCF_3	$MC.sd_1$	$MC.sd_2$	$MC.sd_3$	$asy.sd_1$	$asy.sd_2$	$asy.sd_3$	$boot.sd_1$	$boot.sd_2$	$boot.sd_3$
						cut-point = (2,4)						
True	0.5000	0.4347	0.9347									
FI	0.5011	0.4345	0.9351	0.0277	0.0238	0.0152	0.0239	0.0203	0.0262	0.0275	0.0241	0.0153
MSI	0.5011	0.4344	0.9351	0.0280	0.0259	0.0169	0.0243	0.0226	0.0271	0.0279	0.0260	0.0169
IPW	**0.5822**	**0.4436**	0.9375	0.0381	0.0407	0.0213	0.0380	0.0400	0.0255	0.0381	0.0401	0.0212
SPE	0.5011	0.4345	0.9352	0.0304	0.0334	0.0218	0.0304	0.0330	0.0214	0.0305	0.0331	0.0216
						cut-point = (2,5)						
True	0.5000	0.7099	0.7752									
FI	0.5011	0.7105	0.7765	0.0277	0.0227	0.0297	0.0239	0.0223	0.0334	0.0275	0.0228	0.0298
MSI	0.5011	0.7101	0.7762	0.0280	0.0245	0.0305	0.0243	0.0241	0.0343	0.0279	0.0245	0.0309
IPW	**0.5822**	**0.6815**	**0.8046**	0.0381	0.0376	0.0325	0.0380	0.0370	0.0381	0.0381	0.0371	0.0327
SPE	0.5011	0.7099	0.7760	0.0304	0.0309	0.0328	0.0304	0.0306	0.0330	0.0305	0.0307	0.0331
						cut-point = (2,7)						
True	0.5000	0.9230	0.2248									
FI	0.5011	0.9233	0.2256	0.0277	0.0144	0.0270	0.0239	0.0193	0.0250	0.0275	0.0143	0.0270
MSI	0.5011	0.9234	0.2258	0.0280	0.0161	0.0275	0.0243	0.0204	0.0256	0.0279	0.0158	0.0275
IPW	**0.5822**	**0.9009**	0.2306	0.0381	0.0276	0.0306	0.0380	0.0268	0.0316	0.0381	0.0270	0.0308
SPE	0.5011	0.9234	0.2258	0.0304	0.0225	0.0279	0.0304	0.0218	0.0280	0.0305	0.0220	0.0281
						cut-point = (4,5)						
True	0.9347	0.2752	0.7752									
FI	0.9352	0.2760	0.7765	0.0135	0.0218	0.0297	0.0127	0.0168	0.0334	0.0135	0.0215	0.0298
MSI	0.9352	0.2757	0.7762	0.0143	0.0237	0.0305	0.0135	0.0191	0.0343	0.0143	0.0233	0.0309
IPW	**0.9540**	**0.2379**	**0.8046**	0.0139	0.0335	0.0325	0.0138	0.0330	0.0381	0.0139	0.0331	0.0327
SPE	0.9352	0.2754	0.7760	0.0161	0.0279	0.0328	0.0160	0.0275	0.0330	0.0161	0.0275	0.0331
						cut-point = (4,7)						
True	0.9347	0.4883	0.2248									
FI	0.9352	0.4888	0.2256	0.0135	0.0290	0.0270	0.0127	0.0259	0.0250	0.0135	0.0287	0.0270
MSI	0.9352	0.4889	0.2258	0.0143	0.0302	0.0275	0.0135	0.0273	0.0256	0.0143	0.0300	0.0275
IPW	**0.9540**	**0.4574**	0.2306	0.0139	0.0391	0.0306	0.0138	0.0387	0.0316	0.0139	0.0388	0.0308
SPE	0.9352	0.4890	0.2258	0.0161	0.0328	0.0279	0.0160	0.0327	0.0280	0.0161	0.0328	0.0281
						cut-point = (5,7)						
True	0.9883	0.2132	0.2248									
FI	0.9883	0.2128	0.2256	0.0040	0.0216	0.0270	0.0038	0.0190	0.0250	0.0040	0.0215	0.0270
MSI	0.9884	0.2133	0.2258	0.0050	0.0231	0.0275	0.0046	0.0208	0.0256	0.0050	0.0231	0.0275
IPW	**0.9912**	0.2195	0.2306	0.0060	0.0305	0.0306	0.0054	0.0301	0.0316	0.0059	0.0302	0.0308
SPE	0.9885	0.2135	0.2258	0.0065	0.0256	0.0279	0.0060	0.0256	0.0280	0.0064	0.0257	0.0281

TABLE 16. *Simulation results from 5000 replications when the model for the verification process is misspecified (Study 2) and the third value of* Λ *is used. "True" indicates the true parameter value. Sample size = 1000.*

	TCF_1	TCF_2	TCF_3	$MC.sd_1$	$MC.sd_2$	$MC.sd_3$	$asy.sd_1$	$asy.sd_2$	$asy.sd_3$	$boot.sd_1$	$boot.sd_2$	$boot.sd_3$
						cut-point = (2,4)						
True	0.5000	0.3031	0.8031									
FI	0.4998	0.3026	0.8043	0.0257	0.0172	0.0280	0.0221	0.0124	0.0293	0.0259	0.0171	0.0278
MSI	0.4999	0.3027	0.8044	0.0264	0.0204	0.0297	0.0230	0.0166	0.0308	0.0267	0.0204	0.0293
IPW	**0.6267**	**0.2614**	**0.8259**	0.0345	0.0371	0.0371	0.0346	0.0364	0.0372	0.0348	0.0366	0.0365
SPE	0.5000	0.3031	0.8047	0.0322	0.0323	0.0361	0.0324	0.0321	0.0352	0.0326	0.0322	0.0354
						cut-point = (2,5)						
True	0.5000	0.4682	0.6651									
FI	0.4998	0.4681	0.6667	0.0257	0.0192	0.0341	0.0221	0.0151	0.0342	0.0259	0.0192	0.0343
MSI	0.4999	0.4681	0.6664	0.0264	0.0227	0.0354	0.0230	0.0195	0.0357	0.0267	0.0229	0.0358
IPW	**0.6267**	**0.3884**	**0.7253**	0.0345	0.0403	0.0396	0.0346	0.0400	0.0413	0.0348	0.0402	0.0401
SPE	0.5000	0.4684	0.6665	0.0322	0.0352	0.0389	0.0324	0.0353	0.0391	0.0326	0.0355	0.0393
						cut-point = (2,7)						
True	0.5000	0.7027	0.3349									
FI	0.4998	0.7035	0.3360	0.0257	0.0201	0.0318	0.0221	0.0184	0.0311	0.0259	0.0203	0.0320
MSI	0.4999	0.7035	0.3360	0.0264	0.0237	0.0337	0.0230	0.0224	0.0331	0.0267	0.0240	0.0339
IPW	**0.6267**	**0.6157**	**0.4102**	0.0345	0.0417	0.0386	0.0346	0.0416	0.0398	0.0348	0.0417	0.0386
SPE	0.5000	0.7038	0.3360	0.0322	0.0360	0.0350	0.0324	0.0361	0.0350	0.0326	0.0364	0.0352
						cut-point = (4,5)						
True	0.8031	0.1651	0.6651									
FI	0.8032	0.1655	0.6667	0.0207	0.0139	0.0341	0.0189	0.0099	0.0342	0.0207	0.0141	0.0343
MSI	0.8031	0.1654	0.6664	0.0217	0.0165	0.0354	0.0200	0.0135	0.0357	0.0216	0.0169	0.0358
IPW	**0.8512**	**0.1270**	**0.7253**	0.0217	0.0245	0.0396	0.0215	0.0251	0.0413	0.0215	0.0253	0.0401
SPE	0.8030	0.1653	0.6665	0.0239	0.0225	0.0389	0.0237	0.0228	0.0391	0.0238	0.0229	0.0393
						cut-point = (4,7)						
True	0.8031	0.3996	0.3349									
FI	0.8032	0.4009	0.3360	0.0207	0.0226	0.0318	0.0189	0.0194	0.0311	0.0207	0.0227	0.0320
MSI	0.8031	0.4008	0.3360	0.0217	0.0261	0.0337	0.0200	0.0234	0.0331	0.0216	0.0262	0.0339
IPW	**0.8512**	**0.3544**	**0.4102**	0.0217	0.0358	0.0386	0.0215	0.0362	0.0398	0.0215	0.0363	0.0386
SPE	0.8030	0.4008	0.3360	0.0239	0.0326	0.0350	0.0237	0.0325	0.0350	0.0238	0.0327	0.0352
						cut-point = (5,7)						
True	0.8996	0.2345	0.3349									
FI	0.8997	0.2354	0.3360	0.0144	0.0183	0.0318	0.0135	0.0156	0.0311	0.0144	0.0184	0.0320
MSI	0.8995	0.2354	0.3360	0.0158	0.0223	0.0337	0.0149	0.0197	0.0331	0.0157	0.0220	0.0339
IPW	**0.9149**	**0.2274**	**0.4102**	0.0163	0.0303	0.0386	0.0160	0.0299	0.0398	0.0161	0.0301	0.0386
SPE	0.8995	0.2355	0.3360	0.0175	0.0273	0.0350	0.0173	0.0268	0.0350	0.0174	0.0269	0.0352

TABLE 17
Simulation results for VUS estimators, $\mu = 0.9472$.

Sample size		Mean	MC.sd	Asy.sd	Boot.sd
$n = 200$	FI	0.9471	0.0251	0.0219	0.0256
	MSI	0.9466	0.0252	0.0222	0.0258
	IPW	0.9498	0.0377	0.0261	0.0271
	SPE	0.9461	0.0323	0.0274	0.0315
$n = 500$	FI	0.9470	0.0144	0.0143	0.0149
	MSI	0.9468	0.0144	0.0144	0.0150
	IPW	0.9480	0.0244	0.0192	0.0192
	SPE	0.9467	0.0228	0.0181	0.0224

Appendix C: Simulation results for VUS estimators

In this section, we give some simulation results concerning the estimators of the VUS presented in Subsection 3.6.

The disease $\mathcal{D}$ is generated by a trinomial random vector (D_1, D_2, D_3), such that D_k is a Bernoulli random variable with mean θ_k, $k = 1, 2, 3$. We set $\theta_1 = 0.4, \theta_2 = 0.35$ and $\theta_3 = 0.25$. The pairs T, A are generated from the following conditional models

$$T, A|D_k \sim \mathcal{N}_2\left(\mu_k, \Lambda\right), \qquad k = 1, 2, 3,$$

where $\mu_k = k(\mu_T, \mu_A)^\top$. We consider three values of Λ,

$$\begin{pmatrix} 1.2 & 1 \\ 1 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1.75 & 0.1 \\ 0.1 & 2.5 \end{pmatrix}, \qquad \begin{pmatrix} 5.5 & 3 \\ 3 & 2.5 \end{pmatrix}.$$

The true VUS value is equal to 0.9472 for the first value of Λ and $(\mu_T, \mu_A) = (3, 2)$; is equal to 0.7175 for the second value of Λ and $(\mu_T, \mu_A) = (2, 1)$; is equal to 0.4778 for the third value of Λ and $(\mu_T, \mu_A) = (2, 1)$. We simulate the verification status V by using the following model

$$\text{logit}\left\{\Pr(V = 1|T, A)\right\} = \delta_0 + \delta_1 T + \delta_2 A.$$

The parameters $(\delta_0, \delta_1, \delta_2)$ are fixed equal to $(1, -2.87, 4.06)$ when the first value of Λ is considered, and equal to $(1, -2.2, 4)$ otherwise. These choices give rise to a verification rate of about 0.52. Under our data–generating setting, the disease process follows a multinomial logistic model. We consider two sample sizes, i.e., $n = 200$ and $n = 500$. Each simulation experiment was based on 1000 replications.

FI, MSI, IPW and SPE estimates of VUS are computed under correct working models for both the disease and the verification processes. Tables 17–19 show Monte Carlo means, Monte Carlo standard deviations (MC.sd), the square roots of the variances estimated via asymptotic results (Asy.sd) and bootstrap standard deviations (Boot.sd) of $\hat{\mu}$.

TABLE 18
Simulation results for VUS estimators, $\mu = 0.7175$.

Sample size		Mean	MC.sd	Asy.sd	Boot.sd
$n = 200$	FI	0.7185	0.0549	0.0559	0.0566
	MSI	0.7165	0.0552	0.0571	0.0577
	IPW	0.7261	0.0981	0.1197	0.0754
	SPE	0.7155	0.1021	0.0981	0.1106
$n = 500$	FI	0.7183	0.0357	0.0356	0.0357
	MSI	0.7176	0.0358	0.0360	0.0361
	IPW	0.7272	0.0814	0.0549	0.0564
	SPE	0.7184	0.0813	0.0698	0.0864

TABLE 19
Simulation results for VUS estimators, $\mu = 0.4778$.

Sample size		Mean	MC.sd	Asy.sd	Boot.sd
$n = 200$	FI	0.4788	0.0575	0.0558	0.0574
	MSI	0.4775	0.0584	0.0576	0.0589
	IPW	0.4760	0.1054	0.0767	0.0876
	SPE	0.4815	0.1121	0.1472	0.1418
$n = 500$	FI	0.4782	0.0360	0.0350	0.0354
	MSI	0.4779	0.0364	0.0358	0.0361
	IPW	0.4804	0.0792	0.0608	0.0640
	SPE	0.4868	0.0943	0.1101	0.0995

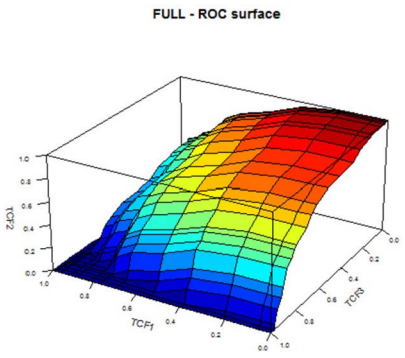


FIG 3. *Estimated ROC surface for CA125 assessing the classification into three class of EOC: benign disease, early stage (I and II) and late stage (I and II). This surface is estimated by using full data.*

Appendix D: Some figures related to the first illustration

In this section, we provide some extra plots related to the analysis of the first dataset used in the main paper. In particular, in Figure 3 we present the estimate of the ROC surface for the test CA125 based on the full data set. Figure 4 and Figure 5 present the projections of the estimated ROC surfaces to the

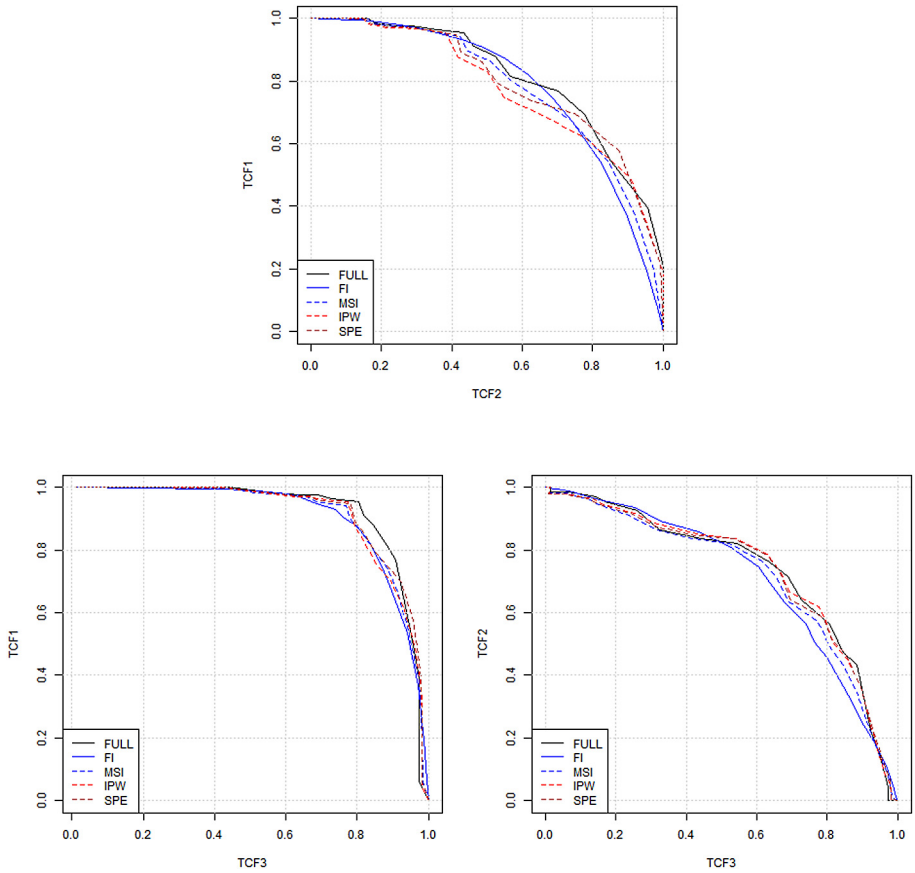


FIG 4. *Two dimensional ROC curve projections. A threshold model is used to estimate the verification process.*

planes defined by TCF_1 versus TCF_2, TCF_1 versus TCF_3 and TCF_2 versus TCF_3, i.e., the ROC curves between classes 1 and 2, classes 1 and 3, classes 2 and 3. For the IPW and SPE methods, to estimate the verification process, we make use, firstly, of a correctly specified model, i.e., a linear threshold regression model (Figure 4) and, then, of a misspecified model, i.e., a logistic model (Figure 5). Finally, as an example, Figure 6 plots confidence regions for the pair $(\mathrm{TCF}_1(c_1), \mathrm{TCF}_2(c_1, +\infty))$ at three values of c_1, when the MSI approach is used. An approximated 95% elliptical confidence region is obtained in a standard way as the set of points

$$R_{12,\mathrm{MSI}} = \left\{ \begin{pmatrix} \mathrm{TCF}_1(c_1) \\ \mathrm{TCF}_2(c_1, +\infty) \end{pmatrix} : \begin{pmatrix} \mathrm{dTCF}_1(c_1) \\ \mathrm{dTCF}_2(c_1, +\infty) \end{pmatrix}^{\top} \hat{\Sigma}_{12}^{-1} \begin{pmatrix} \mathrm{dTCF}_1(c_1) \\ \mathrm{dTCF}_2(c_1, +\infty) \end{pmatrix} \leq \chi^2_{0.95,2}; c_1 \in \mathbb{R} \right\},$$

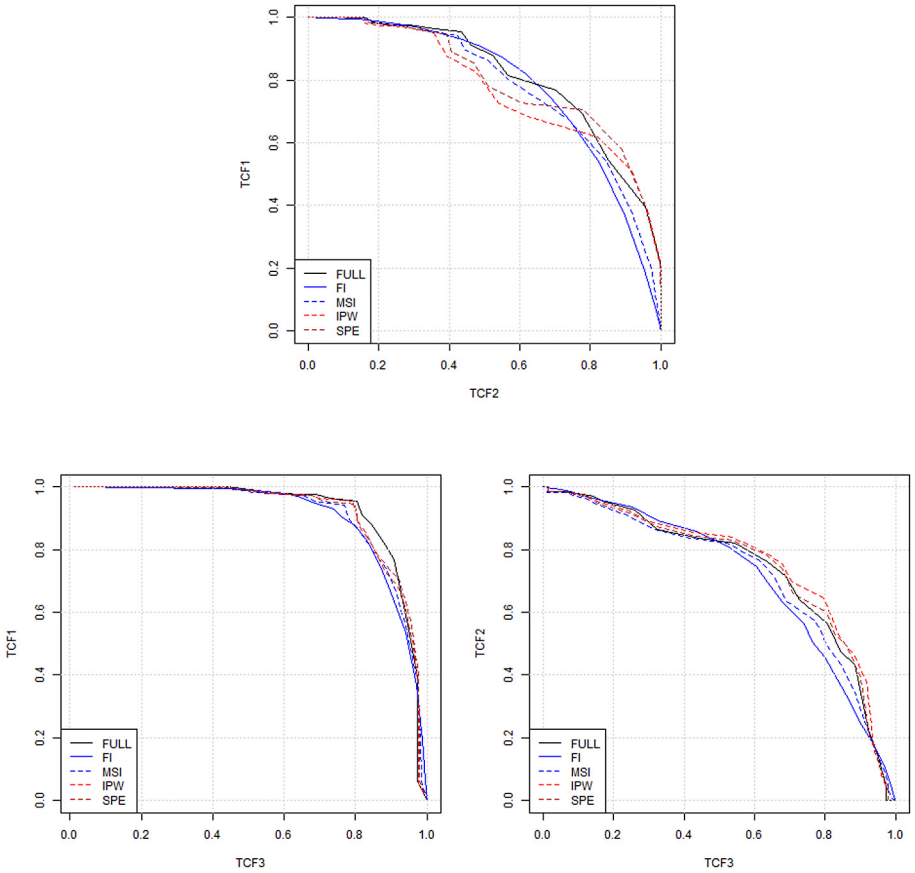


FIG 5. *Two dimensional ROC curve projections. A logistic model is used to estimate the verification process.*

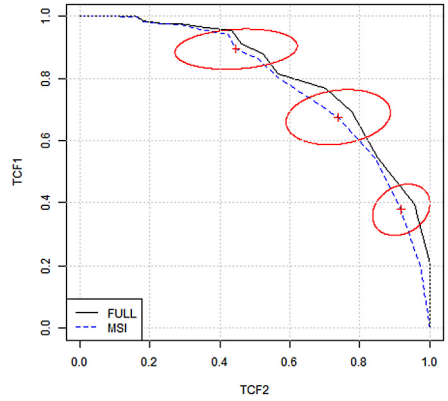


FIG 6. *ROC curve between classes 1 and 2 estimated using the MSI approach along with approximate 0.95 confidence regions for* $c_1 = -0.237, -0.399$ *and 1.672, respectively.*

where

$$\begin{pmatrix} \mathrm{dTCF}_1(c_1) \\ \mathrm{dTCF}_2(c_1,+\infty) \end{pmatrix} = \begin{pmatrix} \mathrm{TCF}_1(c_1) \\ \mathrm{TCF}_2(c_1,+\infty) \end{pmatrix} - \begin{pmatrix} \widehat{\mathrm{TCF}}_{1,\mathrm{MSI}}(c_1) \\ \widehat{\mathrm{TCF}}_{2,\mathrm{MSI}}(c_1,+\infty) \end{pmatrix},$$

the quantity $\hat{\Sigma}_{12}$ is the estimated asymptotic covariance matrix of $\left(\widehat{\mathrm{TCF}}_{1,\mathrm{MSI}}(c_1),\ \widehat{\mathrm{TCF}}_{2,\mathrm{MSI}}(c_1,+\infty)\right)$ and $\chi^2_{0.95,2}$ is the 95–th quantile of a Chi–square distribution with 2 degree of freedom. In the plot, the black solid line represents the full data estimated ROC curve, whereas the blue dashed line is the bias–corrected estimated ROC curve.

References

[1] Adimari, G. and Chiogna, M. (2015). Nearest-neighbor estimation for ROC analysis under verification bias. *The International Journal of Biostatistics* **11** 109–124. MR3341515

[2] Alonzo, T. A., Pepe, M. S. and Lumley, T. (2003). Estimating disease prevalence in two-phase studies. *Biostatistics* **4** 313–326.

[3] Alonzo, T. A. and Pepe, M. S. (2005). Assessing accuracy of a continuous screening test in the presence of verification bias. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **54** 173–190. MR2134605

[4] Chi, Y. Y. and Zhou, X. H. (2008). Receiver operating characteristic surfaces in the presence of verification bias. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **57** 1–23. MR2412663

[5] Daganzo, C. (1979). *Multinomial probit: the theory and its application to demand forecasting*. Elsevier. MR0567139

[6] Fluss, R., Reiser, B., Faraggi, D. and Rotnitzky, A. (2009). Estimation of the ROC curve under verification bias. *Biometrical Journal* **51** 475–490. MR2750048

[7] Gu, J., Ghosal, S. and Kleiner, D. E. (2014). Bayesian ROC curve estimation under verification bias. *Statistics in medicine* **33** 5081–5096. MR3276521

[8] Kang, L. and Tian, L. (2013). Estimation of the volume under the ROC surface with three ordinal diagnostic categories. *Computational Statistics & Data Analysis* **62** 39–51. MR3040261

[9] Li, J. and Zhou, X. H. (2009). Nonparametric and semiparametric estimation of the three way receiver operating characteristic surface. *Journal of Statistical Planning and Inference* **139** 4133–4142. MR2558356

[10] Li, J., Zhou, X. H. and Fine, J. P. (2012). A regression approach to ROC surface, with applications to Alzheimer's disease. *Science China Mathematics* **55** 1583–1595. MR2955244

[11] Little, R. J. and Rubin, D. B. (1987). *Statistical analysis with missing data*. John Wiley & Sons. MR1148365 MR0890519

[12] Nakas, C. T. (2014). Developments in ROC surface analysis and assessment of diagnostic markers in three-class classification problems. *REVSTAT–Statistical Journal* **12** 43–65. MR3195209

[13] Nakas, C. T. and Yiannoutsos, C. T. (2004). Ordered multiple-class ROC analysis with continuous measurements. *Statistics in Medicine* **23** 3437–3449.

[14] Pepe, M. S. (2003). *The statistical evaluation of medical tests for classification and prediction*. Oxford University Press. MR2260483

[15] 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. MR2328313

[16] Scurfield, B. K. (1996). Multiple-event forced-choice tasks in the theory of signal detectability. *Journal of Mathematical Psychology* **40** 253–269.

[17] Xiong, C., van Belle, G., Miller, J. P. and Morris, J. C. (2006). Measuring and estimating diagnostic accuracy when there are three ordinal diagnostic groups. *Statistics in Medicine* **25** 1251–1273. MR2225590

[18] Zhou, X. H., McClish, D. K. and Obuchowski, N. A. (2002). *Statistical methods in diagnostic medicine*. John Wiley & Sons. MR1915698