Optimal weight in estimating and comparing areas under the receiver operating characteristic curve using longitudinal data

Yougui Wu[*,1] and **Xiaofei Wang**[2]

[1] Department of Epidemiology and Biostatistics, College of Public Health, University of South Florida, Tampa, FL 33612, USA
[2] Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC 27710, USA

Received 11 February 2011, revised 26 May 2011, accepted 6 June 2011

In the setting of longitudinal study, subjects are followed for the occurrence of some dichotomous outcome. In many of these studies, some markers are also obtained repeatedly during the study period. Emir et al. introduced a non-parametric approach to the estimation of the area under the ROC curve of a repeated marker. Their non-parametric estimate involves assigning a weight to each subject. There are two weighting schemes suggested in their paper: one for the case when within-patient correlation is low, and the other for the case when within-subject correlation is high. However, it is not clear how to assign weights to marker measurements when within-patient correlation is modest. In this paper, we consider the optimal weights that minimize the variance of the estimate of the area under the ROC curve (AUC) of a repeated marker, as well as the optimal weights that minimize the variance of the AUC difference between two repeated markers. Our results in this paper show that the optimal weights depend not only on the within-patient control–case correlation in the longitudinal data, but also on the proportion of subjects that become cases. More importantly, we show that the loss of efficiency by using the two weighting schemes suggested by Emir et al. instead of our optimal weights can be severe when there is a large within-subject control–case correlation and the proportion of subjects that become cases is small, which is often the case in longitudinal study settings.

Keywords: Asymptotic relative efficiency; Optimal weight; Repeated marker; Sensitivity; Specificity.

Supporting Information for this article is available from the author or on the WWW under http://dx.doi.org/10.1002/bimj.201100033.

1 Introduction

In the setting of longitudinal studies, subjects are followed for the occurrence of some dichotomous outcome. In many of these studies, the markers that might have potential to be used as a surrogate to replace a clinical examination are also obtained repeatedly during the study period. Before these markers can be used with confidence, their accuracy as surrogates of the outcome has to be evaluated relative to the gold standard assessment.

The common approach to the evaluation of the repeated markers is to use repeated markers as an independent variable in a Cox model (Cox, 1972; Kalbfleisch and Prentice, 1980) with outcome as a dependent variable. Emir et al. (1998) presented an example from a breast cancer study to

*Corresponding author: e-mail: ywu@health.usf.edu, Phone: +1-813-974-6673, Fax: +1-813-974-4719

demonstrate why it is more appropriate to use parameters such as specificity and sensitivity than relative hazard that evaluates the relative risk associated with high values of a marker.

Sensitivity is the probability of correctly detecting the disease among subjects with the disease. Specificity is the probability of correctly ruling out the disease among subjects without the disease. DeLong et al. (1985) introduced a method for estimating specificity of a marker in the repeated measurements case when each patient could have at most one event. Their method required direct estimation of the sensitivity and indirect estimation of the specificity through the estimate of relative hazard from a time-dependent Cox model. Emir et al. (1998) proposed a non-parametric approach to estimate both the sensitivity and specificity of a marker when each patient could have more than one event. Their approach involves estimating patient-specific sensitivities and specificities and taking weighted averages of these estimates over all patients.

An ROC curve is a plot of a marker's sensitivity versus 1−specificity. The curve is constructed by changing the cutpoint that defines a positive test result. The area under the ROC curve (AUC) summarizes the marker's overall diagnostic ability and is typically used as a global measure of the accuracy of the marker. In the repeated measurements case when each patient could have at most one event, Parker and DeLong (2003) extended the semi-parametric method of estimating specificity proposed by DeLong et al. (1985) to estimate the AUC. Emir et al. (2000) considered the estimation of the AUC of a marker by generalizing the non-parametric method of estimating specificity proposed by Emir et al. (1998).

The non-parametric estimate of the AUC of a repeated marker by Emir et al. (2000) involves assigning a weight to each subject. There are two weighting schemes suggested in their paper: (i) assigning equal weights to marker measurements when within-patient correlation is low and (ii) assigning equal weights to subjects, when within-subject correlation is high. Intuitively, these two weighting schemes are optimal as each marker measurement is independent in the first case and there is only one effective marker measurement per patient in the second case.

However, it is not clear how to assign weights to marker measurements when within-patient correlation is modest. In this paper, we consider the optimal weights that minimize the variance of the estimate of the area under the ROC curve of a repeated marker proposed by Emir et al. (2000), as well as the optimal weights that minimize the variance of the AUC difference between two repeated markers. Our results in this paper show that the optimal weights depend not only on the within-patient control–case correlation in the longitudinal data, but also on the proportion of subjects that become cases. More importantly, we show that the loss of efficiency by using the two weighting schemes suggested by Emir et al. (2000) instead of our optimal weights can be severe when there is a large within-subject control–case correlation and the proportion of subjects who become cases is small, which is often the case in longitudinal study settings.

The remainder of this paper is organized as follows. In Section 2, the optimal weights for one AUC are derived and the estimators of the optimal weights are discussed. The relative asymptotic efficiencies in comparing our optimal estimator with the two weighting schemes suggested by Emir et al. (2000) are studied. In Section 3, the optimal weights for the difference of two AUCs are derived and the relative asymptotic efficiencies in comparing our optimal estimator with the two weighting schemes suggested by Emir et al. (2000) are studied. In Section 4, a simulation study is conducted to evaluate the finite sample performance of our optimal estimator. A data example is presented in Section 5 and conclusions are provided in Section 6.

2 Optimal weights for one AUC

2.1 Optimal weights derivation

Suppose that we have a random sample of n patients who are being followed for the outcome of disease progression. Let X_{jk} be the continuous random variable whose observations are the marker values obtained from the j-th subject, $j = 1, 2, \ldots, n$, at the k-th non-progression evaluation, $k = 1, 2, \ldots, m_j$, where m_j is the number of non-progression evaluations for subject j. Let $\delta_j = 1$ if subject j became a case and $= 0$ otherwise. Let Y_j be the continuous random variable associated with values for the same marker from the j-th subject for whom $\delta_j = 1$. Also let $D = \sum_{j=1}^{n} \delta_j$ be the total number of cases.

As in Emir et al. (2000), we make the time-independent AUC assumption that $\{X_{jk}, j = 1, \ldots, n,\ k = 1, \ldots, m_j\}$ are marginally identically distributed with CDF F and Y_j have CDF G. Also assume that if the value of X_{jk} or Y_j exceeds a predetermined cut-off point c the marker will be considered positive. Then, the AUC of the marker is $\theta = \int_0^\infty F(t)\,\mathrm{d}G(t)$. Emir et al. (2000) proposed a non-parametric estimate of θ, given by

$$\hat{\theta} = \int_0^\infty \hat{F}(t)\,\mathrm{d}\hat{G}(t), \tag{1}$$

where

$$\hat{F}(t) = \sum_{j=1}^{n} w_j \left\{ \frac{1}{m_j} \sum_{k=1}^{m_j} I(X_{jk} \le t) \right\}, \tag{2}$$

$$\hat{G}(t) = \frac{1}{D} \sum_{j=1}^{n} \delta_j I(Y_j \le t) \tag{3}$$

and $(w_1, \ldots, w_n)$ is a set of weights assigned to subjects satisfying $w_j > 0$, $j = 1, \ldots, n$ and $\sum_{j=1}^{n} w_j = 1$. Two simple weighting schemes were considered in Emir et al. (2000): (i) assigning equal weights to observations, i.e. $w_j = m_j / \sum_{j'=1}^{n} m_{j'}$, when within-subject correlation is low and (ii) assigning equal weights to subjects, i.e. $w_j = 1/n$, when within-subject correlation is high. In the remainder of this section, we derive an optimal weighting scheme and compare it with the two simple weighting schemes proposed by Emir et al. (2000).

To derive our optimal weights, we utilize the following fact:

$$\hat{\theta} - \theta = \sum_{j=1}^{n} (\varepsilon_j + \xi_j) + o_p(n^{-1/2}), \tag{4}$$

where

$$\varepsilon_j = \frac{\delta_j}{D} \int_0^\infty F(t)\,\mathrm{d}\{I(Y_j \le t) - G(t)\},$$
$$\xi_j = \frac{w_j}{m_j} \sum_{k=1}^{m_j} \int_0^\infty \{I(X_{jk} \le t) - F(t)\}\,\mathrm{d}G(t).$$

The proof of the above fact can be found in the Appendix of Emir et al. (2000). Hence, the variance of $\hat{\theta}$ is approximately

$$\mathrm{Var}(\hat{\theta}) = \sum_{j=1}^{n} \mathrm{Var}(\varepsilon_j + \xi_j). \tag{5}$$

Note that

$$\varepsilon_j = \frac{\delta_j}{D} [F(Y_j) - \theta],$$
$$\xi_j = w_j(1 - \theta) - \frac{w_j}{m_j} \sum_{k=1}^{m_j} G(X_{jk}).$$

Defining the transformation

$$U_{jk} = G(X_{jk}), \quad V_j = F(Y_j) \tag{6}$$

we can express the variance of $\hat{\theta}$ in (5) in terms of U_{jk} and V_j as

$$\mathrm{Var}(\hat{\theta}) = \sum_{j=1}^{n} \mathrm{Var}\left\{ w_j \frac{1}{m_j} \sum_{k=1}^{m_j} U_{jk} - \frac{\delta_j}{D} V_j \right\} \tag{7}$$

$$= \sum_{j=1}^{n} (a_j w_j^2 - 2 b_j w_j) + \frac{\sigma_v^2}{D}, \tag{8}$$

where

$$a_j = \mathrm{Var}\left\{ \frac{1}{m_j} \sum_{k=1}^{m_j} U_{jk} \right\} = \frac{\sum_{k,k'=1}^{m_j} \sigma_{kk'}^{uu}}{m_j^2},$$
$$b_j = \frac{\delta_j}{D} \mathrm{Cov}\left\{ \frac{1}{m_j} \sum_{k=1}^{m_j} U_{jk}, V_{1j} \right\} = \frac{\delta_j \sum_{k=1}^{m_j} \sigma_k^{uv}}{D m_j}$$

and $\sigma_{kk'}^{uu} = \mathrm{Cov}(U_{jk}, U_{jk'}), \sigma_{kk}^{uu} = \mathrm{Var}(U_{jk}) = \sigma_u^2$, $\sigma_k^{uv} = \mathrm{Cov}(U_{jk}, V_j)$, and $\sigma_v^2 = \mathrm{Var}(V_j)$.

The optimal weights can be obtained by minimizing (8) with respect to $w = (w_1, \ldots, w_n)$ with constraints $w_j > 0$, $j = 1, \ldots, n$ and $\sum_{j=1}^{n} w_j = 1$. Applying Lagrange multiplier method, we have for $j = 1, \ldots, n$

$$w_j = \frac{1 - \sum_{l=1}^{n} b_l a_l^{-1}}{a_j \sum_{l=1}^{n} a_l^{-1}} + \frac{b_j}{a_j}. \tag{9}$$

The derivation of (9) is given in the Appendix.

The optimal weights involve unknown parameters that need to be estimated. Since they are the variances and covariances with respect to U_{jk}, V_j, we first obtain the estimated transformed data $(\hat{U}_{jk}, \hat{V}_j)$ by

$$\hat{U}_{jk} = \hat{G}(X_{jk}), \quad \hat{V}_j = \hat{F}(Y_j)$$

and then use the estimated transformed data to obtain the estimates of those unknown parameters as given by

$$\hat{\sigma}_{kk'}^{uu} = \frac{\sum_{j=1}^{n} \delta_{jk} \delta_{jk'} (\hat{U}_{jk} - \bar{U}_k)(\hat{U}_{jk'} - \bar{U}_{k'})}{\sum_{j=1}^{n} \delta_{jk} \delta_{jk'}}, \tag{10}$$

$$\hat{\sigma}_k^{uv} = \frac{\sum_{j=1}^{n} \delta_{jk} \delta_j (\hat{U}_{jk} - \bar{U}_k)(\hat{V}_j - \bar{V})}{\sum_{j=1}^{n} \delta_{jk} \delta_j}, \tag{11}$$

where $\delta_{jk} = 1$ if the j-th subject has the k-th non-progression visit and $= 0$ otherwise, $\bar{U}_k = \sum_{j=1}^{n} \delta_{jk} \hat{U}_{jk} / \sum_{j=1}^{n} \delta_{jk}$ and $\bar{V} = \sum_{j=1}^{n} \delta_j \hat{V}_j / D$. The estimated optimal weights are then obtained by plugging those estimates into (9), i.e.

$$\hat{w}_j = \frac{1 - \sum_{k=1}^{n} \hat{b}_k \hat{a}_k^{-1}}{\hat{a}_j \sum_{k=1}^{n} \hat{a}_k^{-1}} + \frac{\hat{b}_j}{\hat{a}_j}, \tag{12}$$

where

$$\hat{a}_j = \frac{\sum_{k,k'=1}^{m_j} \hat{\sigma}_{kk'}^{uu}}{m_j^2},$$
$$\hat{b}_j = \frac{\delta_j \sum_{k=1}^{m_j} \hat{\sigma}_k^{uv}}{Dm_j}.$$

Since $\hat{F}(Y_j)$ includes w_j, we may alternately estimate w_j and $F(\cdot)$ until it converges, or, to allow for a closed-form solution, we may replace $\hat{w}_j$ in $\hat{F}(Y_j)$ with any simple weight such as $w_j = m_j/\sum_{j'=1}^{n} m_{j'}$ or $w_j = 1/n$.

Now consider two special cases: (i) there is no correlation between marker's transformed values at any two time points; (ii) there is a perfect correlation between marker's transformed values at any two time points. In the first case, we have $a_i = \sigma_u^2/m_i$ and $b_i = 0$ and the optimal weight becomes $w_i = m_i/\sum_{i'=1}^{n} m_{i'}$, which means that the simple weighting scheme 1 suggested by Emir et al. (2000) is optimal. In the second case, we have $a_i = \sigma_u^2$ and $b_i = D^{-1}\sigma_u\sigma_v\delta_i$, and the optimal weight becomes $w_i = n^{-1}(1-\sigma_u^{-1}\sigma_v)+D^{-1}\sigma_u^{-1}\sigma_v$ for a subject who became a case at some time point, and $w_i = n^{-1}(1-\sigma_u^{-1}\sigma_v)$ for a subject who remained a control until the end of the study, which means that the simple weighting scheme 2 suggested by Emir et al. (2000) is not optimal unless $\sigma_u = \sigma_v$ and $D = n$, i.e. all subjects became cases at some time point during the study period. The second case illustrates the key feature of our optimal weighting scheme: it assigns more weight to a subject who became a case at some time point than a subject who remains a control at the end of the study, which is a main difference between our optimal weighting scheme and the two simple weighting schemes. As will be seen at the end of this section, the proportion of subjects who become cases plays a significant role in asymptotic relative efficiency comparing our optimal weighting scheme with the two simple weighting schemes

2.2 Asymptotic relative efficiency

Let $\hat{\theta}_{op}$ be the estimate of θ obtained by replacing w_i in (1) by the estimated optimal weights (12). We can show that $\sqrt{n}(\hat{\theta}_{op}-\theta)$ is approximately normal with mean 0 and variance

$$\sigma_{op}^2 = \frac{\left(1 - \mathrm{E}\dfrac{m\sum_{k=1}^{m}\sigma_k^{uv}}{\sum\limits_{k,k'=1}^{m}\sigma_{kk'}^{uu}}\right)^2}{\mathrm{E}\dfrac{m^2}{\sum_{k,k'=1}^{m}\sigma_{kk'}^{uu}}} - \psi^{-1}\mathrm{E}\frac{\left(\sum_{k=1}^{m}\sigma_k^{uv}\right)^2}{\sum_{k,k'=1}^{m}\sigma_{kk'}^{uu}} + \sigma_v^2\psi^{-1}, \tag{13}$$

where $\psi = \lim_{n\to\infty} D/n$, and the expectations are taken with respect to the random variable m, the number of non-progression visits.

Let $\hat{\theta}_1$ be the estimator of θ using the simple weighting scheme 1: $w_j = m_j/\sum_{j'=1}^{n} m_{j'}$, and $\hat{\theta}_2$ be the estimator of θ using the simple weighting scheme 2: $w_j = 1/n$.

To compare our optimal estimator with $\hat{\theta}_1$ and $\hat{\theta}_2$, we consider the asymptotic variances of $\hat{\theta}_1$ and $\hat{\theta}_2$. We can show that $\sqrt{n}(\hat{\theta}_1-\theta)$ is approximately normal $N(0,\sigma_1^2)$, where

$$\sigma_1^2 = \frac{\mathrm{E}\sum_{k,k'=1}^{m}\sigma_{kk'}^{uu}}{(\mathrm{E}m)^2} - 2\frac{\mathrm{E}\sum_{k=1}^{m}\sigma_k^{uv}}{\mathrm{E}m} + \sigma_v^2\psi^{-1} \tag{14}$$

and that $\sqrt{n}(\hat{\theta}_2-\theta)$ is approximately normal $N(0,\sigma_2^2)$, where

$$\sigma_2^2 = \mathrm{E}\frac{\sum_{k,k'=1}^{m}\sigma_{kk'}^{uu}}{m^2} - 2\mathrm{E}\frac{\sum_{k=1}^{m}\sigma_k^{uv}}{m} + \sigma_v^2\psi^{-1}. \tag{15}$$

The proofs of (13), (14) and (15) are included in the Appendix.

Let $\mathrm{RE}_1 = \sigma_{op}^2/\sigma_1^2$ be the asymptotic relative efficiency for comparing $\hat{\theta}_1$ with $\hat{\theta}_{op}$, and $\mathrm{RE}_2 = \sigma_{op}^2/\sigma_2^2$ be the asymptotic relative efficiency for comparing $\hat{\theta}_2$ with $\hat{\theta}_{op}$. In general, both RE_1 and RE_2 depend on $\sigma_{kk'}^{uu}, \sigma_k^{uv}$ and ψ. For illustration purpose, we consider a special case where $\sigma_u^2 = \sigma_v^2$, $\mathrm{Corr}(U_{jk}, U_{jk'}) = \rho_{00}$, and $\mathrm{Corr}(U_{jk}, V_j) = \rho_{01}$. It follows that

$$\mathrm{RE}_1 = \frac{1/\mathrm{E}\frac{m}{1+(m-1)\rho_{00}} - 2\rho_{01} + \psi^{-1} - (\psi^{-1} - 1)\rho_{01}^2 \mathrm{E}\frac{m}{1+(m-1)\rho_{00}}}{\mathrm{E}\left[m(1+(m-1)\rho_{00})\right]/(\mathrm{Em})^2 - 2\rho_{01} + \psi^{-1}}$$

and

$$\mathrm{RE}_2 = \frac{1/\mathrm{E}\frac{m}{1+(m-1)\rho_{00}} - 2\rho_{01} + \psi^{-1} - (\psi^{-1} - 1)\rho_{01}^2 \mathrm{E}\frac{m}{1+(m-1)\rho}}{\mathrm{E}(1+(m-1)\rho_{00})/m - 2\rho_{01} + \psi^{-1}}.$$

To illustrate the effect of ρ_{01}, ρ_{00} and ψ on the asymptotic relative efficiencies, we graph RE_1 and RE_2 against ρ_{01} for three different values of ρ_{00}, $\rho_{00} = 0.5, 0.7, 0.9$ and three different values of ψ, $\psi = 0.3, 0.6, 1$ (Fig. 1). The uniform distribution of m, $\mathrm{P}(m = i) = 1/6$, $i = 1, \ldots, 6$, was used to calculate the expectations involved in RE_1 and RE_2. It can be seen that the efficiency gain of our optimal estimator increases dramatically as ρ_{01} increases and ψ decreases, and increases slowly as ρ_{00} decreases. It shows that the loss of efficiency by using the two weighting schemes suggested in Emir et al. (2000) instead of our optimal weights can be severe when within-subject control–case correlation is large and the proportion of subjects that become cases is small.

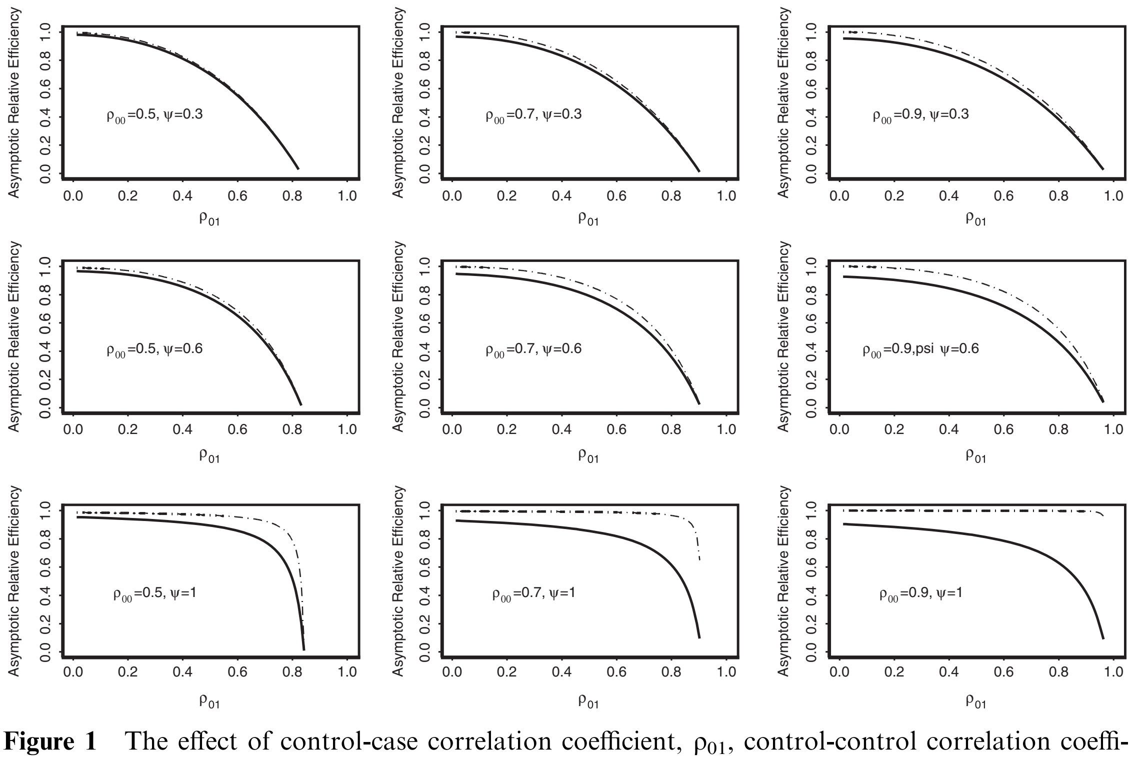

Figure 1 The effect of control-case correlation coefficient, ρ_{01}, control-control correlation coefficient, ρ_{00}, and case proportion, ψ on the asymptotic relative efficiencies, RE_1 (solid line) and RE_2 (broken line).

3 Optimal weights for comparing two AUCs

3.1 Optimal weights derivation

Let X_{ijk} be the continuous random variable whose observations are the i-th marker values obtained from the j-th subject at the k-th non-progression evaluation, $i = 1,2$, $k = 1, 2, \ldots, m_j$, $j = 1, \ldots, n$, where m_j is the number of non-progression evaluations for subject j. Similarly, let Y_{ij} be the continuous random variable associated with values for the i-th marker from the j-th subject at the progression visit. Assume that for each i, $\{X_{ijk}, k = 1, \ldots, m_j, j = 1, \ldots, n\}$ are marginally identically distributed with CDF F_i and Y_{ij} have CDF G_i.

Let β_i be the true AUC for the i-th marker ($i = 1,2$) and $\Delta = \beta_1 - \beta_2$ be the true AUC difference between two markers. The non-parametric estimate of Δ is given by

$$\hat{\Delta} = \int_0^\infty \hat{F}_1(t)\,\mathrm{d}\hat{G}_1(t) - \int_0^\infty \hat{F}_2(t)\,\mathrm{d}\hat{G}_2(t), \tag{16}$$

where

$$\hat{F}_i(t) = \sum_{j=1}^{n} w_j^* \left\{ \frac{1}{m_j} \sum_{k=1}^{m_j} I(X_{ijk} \le t) \right\} \tag{17}$$

$$\hat{G}_i(t) = \frac{1}{D} \sum_{j=1}^{n} \delta_j I(Y_{ij} \le t). \tag{18}$$

We derive the optimal weights to minimize the variance of $\hat{\Delta}$. Similar to the optimal weights for one AUC, we utilize the following fact

$$\hat{\Delta} - \Delta = \sum_{j=1}^{n} (\varepsilon_{1j} + \xi_{1j} - \varepsilon_{2j} - \xi_{2j}) + o(n^{-1/2}),$$

where

$$\varepsilon_{ij} = \frac{\delta_j}{D} \int_0^\infty F_i(t)\,\mathrm{d}\{I(Y_{ij} \le t) - G_i(t)\} = \frac{\delta_j}{D}[F_i(Y_{ij}) - \theta_i]$$

$$\xi_{ij} = \frac{w_j^*}{m_j} \sum_{k=1}^{m_j} \int_0^\infty \{I(X_{ijk} \le t) - F_i(t)\}\,\mathrm{d}G_i(t) = w_j^*(1 - \theta_i) - \frac{w_j^*}{m_j} \sum_{k=1}^{m_j} G_i(X_{ijk}).$$

Defining the transformations

$$U_{ijk} = G_i(X_{ijk}), \quad V_{ij} = F_i(Y_{ij}) \tag{19}$$

and let $U_{jk}^* = U_{1jk} - U_{2jk}$ and $V_j^* = V_{1j} - V_{2j}$, we can express the variance of the estimator of $\hat{\Delta}$ as

$$\begin{aligned} \mathrm{Var}(\hat{\Delta}) &= \sum_{j=1}^{n} \mathrm{Var}(\varepsilon_{1j} - \varepsilon_{2j} + \xi_{1j} - \xi_{2j}) \\ &= \sum_{j=1}^{n} \mathrm{Var}\left\{ w_j^* \frac{1}{m_j} \sum_{k=1}^{m_j} U_{jk}^* - \frac{\delta_j}{D} V_j^* \right\}. \end{aligned} \tag{20}$$

Note that (20) is exactly same as (7) if (U_{jk}, V_j) is replaced by (U_{jk}^*, V_j^*). Therefore, the optimal weights for comparing two AUCs can be obtained by

$$w_j^* = \frac{1 - \sum_{l=1}^{n} b_l^*/a_l^*}{a_j^* \sum_{l=1}^{n} 1/a_l^*} + \frac{b_j^*}{a_j^*} \tag{21}$$

where

$$a_j^* = \frac{\sum_{k,k'=1}^{m_j} \sigma_{kk'}^{u^*u^*}}{m_j^2}$$

$$b_j^* = \frac{\delta_j \sum_{k=1}^{m_j} \sigma_k^{u^*v^*}}{Dm_j}$$

and $\sigma_{kk'}^{u^*u^*} = \mathrm{Cov}(U_{jk}^*, U_{jk'}^*), \sigma_{kk}^{u^*u^*} = \mathrm{Var}(U_{jk}^*) = \sigma_{u^*}^2,\ \sigma_k^{u^*v^*} = \mathrm{Cov}(U_{jk}^*, V_j^*)$, and $\sigma_{v^*}^2 = \mathrm{Var}(V_j^*)$.

The optimal weights (21) contain unknown variance–covariance parameters. They can be estimated in the same way as those in the case of the optimal weights for one AUC, with $(\hat{U}_{jk}, \hat{V}_j)$ replaced by $(\hat{U}_{jk}^*, \hat{V}_j^*)$ where $\hat{U}_{jk}^* = \hat{G}_1(X_{1jk}) - \hat{G}_2(X_{2jk})$ and $\hat{V}_j^* = \hat{F}_1(Y_{1j}) - \hat{F}_2(Y_{2j})$.

3.2 Asymptotic variance comparison

Let $\hat{\Delta}_{op}$ be the estimate of Δ obtained by replacing w_j^* in (16) by the estimated optimal weights (21), $\hat{\Delta}_1$ be the estimator of Δ using simple weighting scheme 1: $w_j^* = m_j / \sum_{j'=1}^n m_{j'}$, and $\hat{\Delta}_2$ be the estimator of Δ using simple weighting scheme 2: $w_j^* = 1/n$.

Along the same line of the proofs for (13), (14) and (15), we can show that $\sqrt{n}(\hat{\Delta}_{op} - \Delta)$ is approximately normal $N(0, \sigma_{op}^{2*})$, $\sqrt{n}(\hat{\Delta}_1 - \Delta)$ is approximately normal $N(0, \sigma_1^{2*})$, and $\sqrt{n}(\hat{\Delta}_2 - \Delta)$ is approximately normal $N(0, \sigma_2^{2*})$, where

$$\sigma_{op}^{2*} = \frac{\left(1 - \mathrm{E}\frac{m\sum_{k=1}^m \sigma_k^{u^*v^*}}{\sum_{k,k'=1}^m \sigma_{kk'}^{u^*u^*}}\right)^2}{\mathrm{E}\frac{m^2}{\sum_{k,k'=1}^m \sigma_{kk'}^{u^*u^*}}} - \psi^{-1}\mathrm{E}\frac{\left(\sum_{k=1}^m \sigma_k^{u^*v^*}\right)^2}{\sum_{k,k'=1}^m \sigma_{kk'}^{u^*u^*}} + \sigma_{v^*}^2\psi^{-1},$$

$$\sigma_1^{2*} = \frac{\mathrm{E}\sum_{k,k'=1}^m \sigma_{kk'}^{u^*u^*}}{(\mathrm{E}m)^2} - 2\frac{\mathrm{E}\sum_{k=1}^m \sigma_k^{u^*v^*}}{\mathrm{E}m} + \sigma_{v^*}^2\psi^{-1},$$

$$\sigma_2^{2*} = \mathrm{E}\frac{\sum_{k,k'=1}^m \sigma_{kk'}^{u^*u^*}}{m^2} - 2\mathrm{E}\frac{\sum_{k=1}^m \sigma_k^{u^*v^*}}{m} + \sigma_{v^*}^2\psi^{-1}.$$

Let $\mathrm{RE}_1^* = \sigma_{op}^{2*}/\sigma_1^{2*}$ be the asymptotic relative efficiency for comparing $\hat{\Delta}_1$ with $\hat{\Delta}_{op}$, and $\mathrm{RE}_2^* = \sigma_{op}^{2*}/\sigma_2^{2*}$ be the asymptotic relative efficiency for comparing $\hat{\Delta}_2$ with $\hat{\Delta}_{op}$. To see how within-marker, between-marker correlations, and the proportion of subjects becoming cases affect RE_1^* and RE_2^*, we consider a special case where variances of U_{ijk} and V_{ij} are homogeneous across visits and markers, within-marker control–control correlation coefficients are homogeneous across markers, within-marker control–case correlation coefficients are homogeneous across markers, and between-marker correlation coefficients are homogeneous across visits. Specifically,

$$\mathrm{Var}(U_{ijk}) = \mathrm{Var}(V_{ij}) = \sigma^2,$$
$$\mathrm{Corr}(U_{ijk}, U_{ijk'}) = \rho_{00},\ \mathrm{Corr}(U_{ijk}, V_{ij}) = \rho_{01},$$
$$\mathrm{Corr}(U_{ijk}, U_{i'jk}) = \mathrm{Corr}(U_{ijk}, V_{i'j}) = \rho_b, i \neq i'.$$

Let $\rho_{00}^* = \mathrm{Corr}(U_{jk}^*, U_{jk'}^*)$, and $\rho_{01}^* = \mathrm{Corr}(U_{jk}^*, V_j^*)$. Then we have

$$\rho_{00}^* = \frac{\rho_{00} - \rho_b}{1 - \rho_b}, \tag{22}$$

$$\rho_{01}^* = \frac{\rho_{01} - \rho_b}{1 - \rho_b}. \tag{23}$$

Similar to the case of a single AUC, it follows that

$$\mathrm{RE}_1^* = \frac{1/\mathrm{E}\frac{m}{1+(m-1)\rho_{00}^*} - 2\rho_{01}^* + \psi^{-1} - (\psi^{-1}-1)\rho_{01}^{*2}\mathrm{E}\frac{m}{1+(m-1)\rho_{00}^*}}{\mathrm{E}[m(1+(m-1)\rho_{00}^*)]/(\mathrm{E}m)^2 - 2\rho_{01}^* + \psi^{-1}},$$

$$\mathrm{RE}_2^* = \frac{1/\mathrm{E}\frac{m}{1+(m-1)\rho_{00}^*} - 2\rho_{01}^* + \psi^{-1} - (\psi^{-1}-1)\rho_{01}^{*2}\mathrm{E}\frac{m}{1+(m-1)\rho_{00}^*}}{\mathrm{E}(1+(m-1)\rho_{00}^*)/m - 2\rho_{01}^* + \psi^{-1}}$$

and both RE_1^* and RE_2^* increase dramatically as ρ_{01}^* increases. Note that ρ_{01}^* decreases as ρ_b increases. Therefore, we may conclude that the large value of ρ_b leads to the large values of RE_1^* and RE_2^*. The effects of ψ, ρ_{00} and ρ_{01} on RE_1^* and RE_2^* are similar to RE_1 and RE_2.

4 Simulation results

To evaluate the finite sample performance of our optimal estimators in comparison with their counterparts using simple weighting schemes 1 and 2, we performed simulations that are similar to the ones in Emir et al. (2000).

4.1 One marker case

In this simulation, we assume that the marker is obtained every year for a total of at most seven annual visits per subject. For each subject, we generate an independent multivariate normal random vector $\mathbf{X} = (X_1, \ldots, X_6)'$ of size 6 with mean vector 0 and variance–covariance matrix $\mathbf{\Sigma}$ with $\mathrm{Cov}(X_k, X_{k'}) = \gamma^{|k-k'|}$, for $k, k' = 1, \ldots, 6$ and $0 \le \gamma < 1$. We generate failure times for the patients using an exponential distribution such that the expected failure rate at six months is ψ.

If a simulated failure time is greater than six months, we denote the patient to be a control at all six visits and use all six values of $\mathbf{X}$ for the marker values. If a simulated failure time occurs before six months, we assume that the failure is detected clinically at the next visit. For example, if a failure occurs between the third and fourth visit, the simulated markers for the first three visits are (x_1,x_2,x_3). We assume the expected value of the marker is increased by 0.5 at the time of failure, hence we define the marker at this fourth visit to be $Y = x_{14}+0.5$. With this setup, the true AUC of the simulated marker is $\theta = \int_0^1 \{1 - \Psi(\Psi^{-1}(t) - 0.5)\}\,\mathrm{d}t = 0.63816$, where Ψ is the CDF of the standard normal distribution.

In this simulation, we set $n = 100$ with $\gamma = 0, 0.3, 0.5, 0.7, 0.9$ and $\psi = 0.3, 0.5$. For each combination of (γ,ψ), we generated 2000 samples. For each sample, we obtained the three point estimates of θ: $\hat{\theta}_{op}$, $\hat{\theta}_1$ and $\hat{\theta}_2$. Table 1 reports the comparison in terms of the mean, and the mean-squared error (MSE). The conclusion from this simulations is that our optimal estimator outperforms both $\hat{\theta}_1$ and $\hat{\theta}_2$ especially when γ is large and ψ is small, which is consistent with the conclusion from the asymptotic comparison.

4.2 Two markers case

In this simulation, for each subject, we generate an independent multivariate normal random vector $(X_{11}, \ldots, X_{16}, X_{21}, \ldots, X_{26})'$ of size 121 with mean vector $\mathbf{0}$ and variance–covariance matrix $\mathbf{\Sigma} = (\sigma_{kk'})_{12\times 12}$ with

$$\sigma_{kk'} = \begin{cases} 1 & \text{if } k = k', \\ \gamma & \text{if } k,\ k' = 1, \ldots, 6, \quad \text{or} \quad k,\ k' = 7, \ldots, 12, \\ \lambda & \text{if } k = 1, \ldots, 6,\ k' = 7, \ldots, 12, \quad \text{or} \quad k = 7, \ldots, 12,\ k' = 1, \ldots, 6, \end{cases}$$

where γ is the within-marker correlation coefficient and λ is the between-marker correlation coefficient. This correlation structure is different from that by Emir et al. (2000), which is not suited for the current simulation purpose because it leads to the independence of λ on the correlation between two marker differences. We chose this particular correlation structure so that there are a wide range of values for γ and λ for which the variance–covariance matrix $\mathbf{\Sigma}$ is positive definite.

We generate failure times for the patients using an exponential distribution such that the expected failure rate at six months is ψ. If a simulated failure time is greater than six months, we denote the patient to be a control at all six visits and use all six values $(x_{i1}, \ldots, x_{i6})$ for the values of the i-th marker ($i = 1,2$). If a simulated failure time occurs before six months, we assume that the failure is detected clinically at the next visit. For example, if a failure occurs between the third and fourth

Table 1 Simulation results for one AUC. The true AUC value for the simulated marker $\theta = 0.63816$. The sample size $n = 100$.

		Optimal		Weight 1		Weight 2	
ψ	γ	Mean	MSE	Mean	MSE	Mean	MSE
0.3	0.0	0.63790	0.00297	0.63825	**0.00279**	0.63789	0.00284
0.3	0.3	0.64025	0.00266	0.64053	0.00266	0.64048	**0.00258**
0.3	0.5	0.63728	**0.00246**	0.63846	0.00269	0.63829	0.00248
0.3	0.7	0.64004	**0.00194**	0.64054	0.00268	0.64081	0.00231
0.3	0.9	0.64054	**0.00108**	0.64011	0.00246	0.64072	0.00195
0.5	0.0	0.63680	0.00172	0.63693	**0.00167**	0.63681	0.00172
0.5	0.3	0.63842	0.00171	0.63843	0.00174	0.63828	**0.00167**
0.5	0.5	0.63827	**0.00148**	0.63813	0.00173	0.63832	0.00149
0.5	0.7	0.63918	**0.00113**	0.63822	0.00165	0.63867	0.00129
0.5	0.9	0.64121	**0.00060**	0.64076	0.00151	0.64115	0.00098

Table 2 Simulation results for two AUCs. The two true AUC values for the simulated markers are equal ($\beta_1 = \beta_2 = 0.63816$). The sample size $n = 100$ and $\psi = 0.3$.

		Optimal		Weight 1		Weight 2	
γ	λ	Mean	MSE	Mean	MSE	Mean	MSE
0.5	0.0	0.00151	**0.00398**	0.00213	0.00540	0.00212	0.00476
0.5	0.3	−0.00159	**0.00356**	−0.00115	0.00380	−0.00157	0.00360
0.5	0.5	0.00231	0.00303	0.00201	**0.00284**	0.00255	0.00292
0.7	0.0	0.00034	**0.00285**	0.00154	0.00513	0.00144	0.00427
0.7	0.3	0.00025	**0.00258**	−0.00117	0.00373	−0.00083	0.00325
0.7	0.5	−0.00097	**0.00237**	−0.00029	0.00272	−0.00039	0.00250
0.7	0.7	0.00009	0.00173	0.00027	**0.00163**	0.00052	0.00168
0.9	0.0	0.00108	**0.00159**	0.00209	0.00530	0.00196	0.00406
0.9	0.3	−0.00033	**0.00155**	−0.00206	0.00375	−0.00191	0.00295
0.9	0.5	−0.00081	**0.00134**	−0.00022	0.00277	−0.00041	0.00222
0.9	0.7	0.00005	**0.00113**	0.00029	0.00174	0.00030	0.00146
0.9	0.9	−0.00089	0.00070	−0.00056	**0.00066**	−0.00050	0.00068

visits, the i-th simulated marker for the first three visits are (x_{i1}, x_{i2}, x_{i3}), $i = 1,2$. We assume the expected values of the markers are increased by 0.5 at the time of failure, hence we define the i-th marker at this fourth visit to be $x_{i4}+0.5$, $i = 1,2$. With this set-up, the two true AUC values for the simulated markers are equal ($\beta_1 = \beta_2 = 0.63816$) and hence the AUC difference of two markers $\Delta = 0$.

In our simulation, we set $n = 100$, and $\psi = 0.3$. The values of λ and γ were chosen such that the variance–covariance matrix $\mathbf{\Sigma}$ is positive definite. For each combination of (λ,γ), we generated 2000 samples. For each sample, we obtained the three point estimates of Δ: $\hat{\Delta}_{op}$, $\hat{\Delta}_1$ and $\hat{\Delta}_2$. Table 2 reports the comparison in terms of the mean and the MSE. The conclusion from this simulations is that our optimal estimator outperforms both $\hat{\Delta}_1$ and $\hat{\Delta}_2$ when the within-marker correlation is large. In addition, it seems that the performance is negatively affected by the between-marker correlation.

5 Data example

Previous studies have found that elevated levels of vascular enothelial growth factor (VEGF) and a soluble fragment of Cytokeratin 19 (CYFRA) are common in patients with non-small cell lung cancer (NSCLC); higher levels of the two markers correlate with worse survival in these patients. The Cancer and Leukemia Group B (CALGB) conducted a phase II study in advanced NSCLC patients (Edelman et al., 2007) to evaluate the efficacy of carboplatin/gemcitabine with eicosanoid modulators (celecoxib, zileuton or both). The objective of its correlative science study is to prospectively measure VEGF and CYFRA in serum collected from enrolled patients and to determine whether the levels of these markers had prognostic significance to disease progression and which marker is more predictive. Patients with advanced NSCLC were treated with a maximum of 6 cycles (4 weeks/cycle) with the combination therapy. All eligible patients had a lesion which was measurable or evaluable and could be followed for disease progression. These subjects had CT scans for restaging every 2 cycles. The status of disease progression was documented at each scan. The study also requires submission serum prior to treatment, at cycles 2, 4 and 6 and then every 2 months for 2 years. Blood was drawn and sent to a central pathology laboratory for analysis. Serum VEGF levels were determined with a commercially available ELISA assay utilizing colormetric endpoints and with a range of 31–2000 pg/mL. Serum CYFRA levels were measured using an electrochemoluminescent assay (Roche Diagnostics) on the ElecSys 2010 system. The analysis in this section utilize the data for serum VEGF and CYFRA for 80 patients, all of whom had at least a baseline value and one follow-up value for the two markers. Over 35% of the patients were followed to progression and those subjects who are progression-free subjects had at least two years of follow-up. These patients had a total of 196 non-progression visits and 40 progression visits.

Log was taken on the raw value of serum VEGF, which is treated as the original data of (X_{jk}, Y_j). The optimal weights for the AUC of the log-transformed VEGF were obtained by implementing a four-step procedure: step 1: apply simple weight 1 to (2) to obtain the estimate of F and use (3) to obtain the estimate of G; step 2: replace F and G in (6) by the estimates of F and G, respectively to obtain the transformed data $(\hat{U}_{ij}, \hat{V}_j)$; step 3: use (10) to obtain the estimates of $\sigma^{uu}_{kk'}$: $\hat{\sigma}^{uu}_{11} = 0.0684$, $\hat{\sigma}^{uu}_{12} = 0.0489$, $\hat{\sigma}^{uu}_{13} = 0.0618$, $\hat{\sigma}^{uu}_{22} = 0.0818$, $\hat{\sigma}^{uu}_{23} = 0.0578$, and $\hat{\sigma}^{uu}_{33} = 0.0833$, and use (11) to obtain the estimates of σ^{uv}_j: $\hat{\sigma}^{uv}_1 = 0.0416$, $\hat{\sigma}^{uv}_2 = 0.0413$; step 4: use Eq. (12) to obtain the optimal weight estimates. The estimated optimal weights were then plugged into Eq. (1) to calculate the point estimate of the AUC for this marker. Equation (8) was used to calculate the estimate of the variance of $\hat{\theta}_{op}$ using the estimated optimal weights. Both point estimate and variance estimate using the optimal weights were included in Table 3. For comparison purpose, both point estimates and variance estimates using weight 1 and 2 were also included in Table 3. It can be seen that our optimal weight performs better than simple weight 1 and 2 with significant efficiency gain.

Log was taken on the raw value of VEGF and CYFRA, which is treated as the original data of (X_{ijk}, Y_{ij}). The optimal weights for the AUC difference between log(VEGF) and log(CYFRA), were obtained by implementing a similar four-step procedure: step 1: apply simple weight 1 to (17) to

Table 3 Results for the VEGF and CYFRA data.

	Single AUC			Two AUCs difference		
	Optimal	Weight 1	Weight 2	Optimal	Weight 1	Weight 2
Estimate	0.608	0.556	0.570	−0.0202	−0.0515	−0.0303
Variance	0.00218	0.00302	0.00275	0.0034	0.0037	0.0035
Relative efficiency	1	0.72	0.79	1	0.92	0.97

obtain the estimate of F_i and use (18) to obtain the estimate of G_i; step 2: replace F_i and G_i in (19) by the estimates of F_i and G_i respectively to obtain the transformed data $(\hat{U}^*_{jk}, \hat{V}^*_j)$; step 3: use the transformed data $\hat{U}^*_{jk}$ and (10) to obtain the estimates of $\sigma^{u^*u^*}_{kk'}$: $\hat{\sigma}^{u^*u^*}_{11} = 0.147$, $\hat{\sigma}^{u^*u^*}_{12} = 0.0861$, $\hat{\sigma}^{u^*u^*}_{13} = 0.141$, $\hat{\sigma}^{u^*u^*}_{22} = 0.164$, $\hat{\sigma}^{u^*u^*}_{23} = 0.114$, and $\hat{\sigma}^{u^*u^*}_{33} = 0.193$, and use the transformed data $\hat{V}^*_j$ and (11) to obtain the estimates of σ^{uv*}_j: $\hat{\sigma}^{uv*}_1 = 0.0062$, $\hat{\sigma}^{uv*}_2 = 0.0368$; step 4: plug these estimated variance–covariance parameters into (21) to obtain the optimal weight estimates. The estimated optimal weights were then plugged into Eq. (16) to calculate the point estimate of the two AUCs difference, Δ. Equation (20) was used to calculate the estimate of the variance of $\hat{\Delta}_{op}$ using the estimated optimal weights. Both point estimate and variance estimate using the optimal weights are included in Table 3. For comparison purpose, both point estimates and variance estimates using weights 1 and 2 are also included in Table 3. It can be seen that our optimal weight performs better than simple weights 1 and 2 with marginal efficiency gain, which is expected because the controlcase correlation in this data example is not large (about 0.1).

6 Concluding remarks

We have derived the optimal weights that minimize the variance of the estimate of the AUC of a repeated marker, as well as the optimal weights that minimize the variance of the AUC difference two repeated markers. Unlike the two simple weighting schemes, our optimal weighting scheme assigns more weight to a subject who became a case at some point than a subject who remained a control at the end of the study. Both asymptotic and finite sample performance of the AUC estimate using our optimal weights have been studied in contrast with the two simple weighting schemes suggested by Emir et al. (2000). We have shown that when there is a large within-subject control–case correlation and the proportion of subjects that become cases is small, using either of the two weighting schemes suggested by Emir et al. (2000) can lead to dramatic efficiency loss, and our optimal weights are recommended.

Conflict of interest
The authors have declared no conflict of interest.

Appendix

Derivation of (9): Equation (9) is derived by applying Lagrange multiplier method. Let

$$H = \sum_{j=1}^{n}(a_j w_j^2 - 2b_j w_j) + \frac{\sigma_v^2}{D} - \lambda\left(\sum_{j=1}^{n} w_i - 1\right),$$

where λ is a Langrage multiplier. Taking derivatives with respective to w_j, $j = 1, \ldots, n$, we have

$$\frac{\partial H}{\partial w_j} = 2a_j w_j - 2b_j - \lambda = 0,$$

which gives

$$w_j = \frac{b_j}{a_j} + \frac{\lambda}{2a_j}. \tag{A1}$$

Using the constraint $\sum_{j=1}^{n} w_j = 1$, we have

$$\lambda = \frac{2(1 - \sum_{j=1}^{n} \frac{b_j}{a_j})}{\sum_{j=1}^{n} a_j^{-1}}. \tag{A2}$$

Plugging (A2) back into (A1) gives

$$w_j = \frac{1 - \sum_{l=1}^{n} b_l a_l^{-1}}{a_j \sum_{l=1}^{n} a_l^{-1}} + \frac{b_j}{a_j}.$$

Proof of (13): Using the fact (4), to prove (13), it suffices to show that

$$\sigma_{op}^2 = \lim_{n\to\infty} n \operatorname{Var}(\hat{\theta}_{op}).$$

Replacing weights in (8) by the optimal weights (9), we have

$$\begin{aligned} n\operatorname{Var}(\hat{\theta}_{op}) &= n\sum_{j=1}^{n} a_j \left(\frac{1 - \sum_{l=1}^{n} b_l a_l^{-1}}{a_j \sum_{l=1}^{n} a_l^{-1}} + \frac{b_j}{a_j}\right)^2 - 2n\sum_{j=1}^{n} b_j \left(\frac{1 - \sum_{l=1}^{n} b_l a_l^{-1}}{a_j \sum_{l=1}^{n} a_l^{-1}} + \frac{b_j}{a_j}\right) + \sigma_v^2 \frac{n}{D} \\ &= \frac{(1 - \sum_{j=1}^{n} b_j a_j^{-1})^2}{n^{-1} \sum_{j=1}^{n} a_j^{-1}} - n\sum_{j=1}^{n} \frac{b_j^2}{a_j} + \sigma_v^2 \frac{n}{D}. \end{aligned}$$

Note that

$$\begin{aligned} &\lim_{n\to\infty} \sum_{j=1}^{n} b_j a_j^{-1} = \lim_{n\to\infty} \frac{1}{D} \sum_{j=1}^{n} \delta_j \frac{m_j \sum_{k=1}^{m_j} \sigma_k^{uv}}{\sum_{k,k'=1}^{m_j} \sigma_{kk'}^{uu}} = \mathrm{E}\, \frac{m \sum_{k=1}^{m} \sigma_k^{uv}}{\sum_{k,k'=1}^{m} \sigma_{kk'}^{uu}}, \\ &\lim_{n\to\infty} n^{-1} \sum_{j=1}^{n} a_j^{-1} = n^{-1} \sum_{j=1}^{n} \frac{m_j^2}{\sum_{k,k'=1}^{m_j} \sigma_{kk'}^{uu}} = \mathrm{E}\, \frac{m^2}{\sum_{k,k'=1}^{m} \sigma_{kk'}^{uu}} \end{aligned}$$

and

$$\lim_{n\to\infty} n \sum_{j=1}^{n} b_j^2 a_j^{-1} = \lim_{n\to\infty} \frac{n}{D} \lim_{n\to\infty} \frac{1}{D} \sum_{j=1}^{n} \delta_j \frac{\left(\sum_{k=1}^{m_j} \sigma_k\right)^2}{\sum_{k,k'=1}^{m_j} \sigma_{kk'}} = \psi^{-1} \mathrm{E}\, \frac{\left(\sum_{k=1}^{m} \sigma_k^{uv}\right)^2}{\sum_{k,k'=1}^{m} \sigma_{kk'}^{uu}}.$$

Therefore, we have

$$\lim_{n\to\infty} n\operatorname{Var}(\hat{\theta}_{op}) = \frac{\left(1 - \mathrm{E}\, \frac{m \sum_{k=1}^{m} \sigma_k^{uv}}{\sum_{k,k'=1}^{m} \sigma_{kk'}^{uu}}\right)^2}{\mathrm{E}\, \frac{m^2}{\sum_{k,k'=1}^{m} \sigma_{kk'}^{uu}}} - \psi^{-1} \mathrm{E}\, \frac{\left(\sum_{k=1}^{m} \sigma_k^{uv}\right)^2}{\sum_{k,k'=1}^{m} \sigma_{kk'}} + \sigma_v^2 \psi^{-1}.$$

Proof of (14): Replacing weights in (8) by the simple weight 1, we have

$$
\begin{aligned}
n\mathrm{Var}(\hat{\theta}_1) =& n\sum_{j=1}^{n} a_j \left(\frac{m_j}{\sum_{j'=1}^{n} m_{j'}}\right)^2 - 2n\sum_{j=1}^{n} b_j \frac{m_j}{\sum_{j'=1}^{n} m_{j'}} + \sigma_v \frac{n}{D} \\
=& \frac{\frac{1}{n}\sum_{j=1}^{m}\sum_{k,k'=1}^{m_j}\sigma_{kk'}^{uu}}{\left(\frac{1}{n}\sum_{j'=1}^{n} m_{j'}\right)^2} - 2\frac{\frac{1}{D}\sum_{j=1}^{n}\delta_j\sum_{k=1}^{m_j}\sigma_k^{uv}}{\frac{1}{n}\sum_{j'=1}^{n} m_{j'}} + \sigma_v^2\frac{n}{D} + \sigma_v^2\psi^{-1}. \\
\to& \frac{\mathrm{E}\sum_{k,k'=1}^{m}\sigma_{kk'}^{uu}}{(\mathrm{E}m)^2} - 2\frac{\mathrm{E}\sum_{k=1}^{m}\sigma_k^{uv}}{\mathrm{E}m}
\end{aligned}
$$

Proof of (15): Replacing weights in (8) by the simple weight 2, we have

$$
\begin{aligned}
n\mathrm{Var}(\hat{\theta}_2) =& n\sum_{j=1}^{n} a_j \frac{1}{n^2} - 2n\sum_{j=1}^{n} b_j \frac{1}{n} + \sigma_v^2\frac{n}{D} \\
=& \frac{1}{n}\sum_{j=1}^{n}\frac{\sum_{k,k'=1}^{m_j}\sigma_{kk'}^{uu}}{m_j^2} - 2\frac{1}{D}\sum_{j=1}^{n}\delta_j\frac{\sum_{k=1}^{m_j}\sigma_k^{uv}}{m_j} + \sigma_v^2\frac{n}{D} \\
\to& \mathrm{E}\frac{\sum_{k,k'=1}^{m}\sigma_{kk'}^{uu}}{m^2} - 2\mathrm{E}\frac{\sum_{k=1}^{m}\sigma_k^{uv}}{m} + \sigma_v^2\psi^{-1}.
\end{aligned}
$$

Proof of (22):

$$
\begin{aligned}
\rho_{00}^* =& \frac{\mathrm{Cov}(U_{jk}^*, U_{jk'}^*)}{\sqrt{\mathrm{Var}(U_{jk}^*)\mathrm{Var}(U_{jk'}^*)}} = \frac{\mathrm{Cov}(U_{1jk} - U_{2jk})(U_{1jk'} - U_{2jk'})}{\sqrt{\mathrm{Var}(U_{1jk} - U_{2jk})\mathrm{Var}(U_{1jk'} - U_{2jk'})}} \\
=& \frac{\rho_{00}\sigma^2 - \rho_b\sigma^2 - \rho_b\sigma^2 + \rho_{00}\sigma^2}{\sqrt{(2\sigma^2 - 2\rho_b\sigma^2)(2\sigma^2 - 2\rho_b\sigma^2)}} = \frac{\rho_{00} - \rho_b}{1 - \rho_b}.
\end{aligned}
$$

Proof of (23):

$$
\begin{aligned}
\rho_{01}^* =& \frac{\mathrm{Cov}(U_{jk}^*, V_j^*)}{\sqrt{\mathrm{Var}(U_{jk}^*)\mathrm{Var}(V_j^*)}} = \frac{\mathrm{Cov}(U_{1jk} - U_{2jk})(V_{1j} - V_{2j})}{\sqrt{\mathrm{Var}(U_{1jk} - U_{2jk})\mathrm{Var}(V_{1j} - V_{2j})}} \\
=& \frac{\rho_{01}\sigma^2 - \rho_b\sigma^2 - \rho_b\sigma^2 + \rho_{01}\sigma^2}{\sqrt{(2\sigma^2 - 2\rho_b\sigma^2)(2\sigma^2 - 2\rho_b\sigma^2)}} = \frac{\rho_{01} - \rho_b}{1 - \rho_b}.
\end{aligned}
$$

References

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B* **34**, 187–220.

DeLong, E. R., Vernon, W. B. and Bollinger, R. R. (1985). Sensitivity and specificity of a monitoring test. *Biometrics*, **41**, 947–958.

Edelman, M. J., Watson, D., Wang, X. F., Morrison, C., Kratzke, R., Jewell, S., Hodgson, L., Mauer, A. M., Graziano, S. L., Masters, G. A., Bedor, M., Green, M. J. and Vokes, E. E. (2007). Eicosanoid modulation in advanced lung cancer: COX-2 expression is a positive predictive factor for celecoxib+chemotherapy. *Journal of Clinical Oncology* **26**, 848–855.

Emir, B., Wieand, S., Jung, S. and Ying, Z. (2000). Comparison of diagnostic markers with repeated measurements: a non-parametric ROC curve approach. *Statistics in Medicine* **19**, 511–523.

Emir, B., Wieand, S., Su, J. and Cha, S. (1998). Analysis of repeated markers used to predict progression of cancer. *Statistics in Medicine* **17**, 2563–2578.

Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*, Wiley, New York, **34**, 187–220.

Parker, C. B. and DeLong, E. R. (2003). ROC methodology within a monitoring framework. *Statistics in Medicine* **22**, 3473–3488.

