



Published in final edited form as:

Ann Stat. 2011 ; 39(6): 3234–3261. doi:10.1214/11-AOS937.

Asymptotic Properties of the Sequential Empirical ROC, PPV and NPV Curves Under Case-Control Sampling

Joseph S. Koopmeiners^{*,†,‡} and Ziding Feng^{*,†,§}

[‡]University of Minnesota

[§]Fred Hutchinson Cancer Research Center

Abstract

The receiver operating characteristic (ROC) curve, the positive predictive value (PPV) curve and the negative predictive value (NPV) curve are three measures of performance for a continuous diagnostic biomarker. The ROC, PPV and NPV curves are often estimated empirically to avoid assumptions about the distributional form of the biomarkers. Recently, there has been a push to incorporate group sequential methods into the design of diagnostic biomarker studies. A thorough understanding of the asymptotic properties of the sequential empirical ROC, PPV and NPV curves will provide more flexibility when designing group sequential diagnostic biomarker studies. In this paper we derive asymptotic theory for the sequential empirical ROC, PPV and NPV curves under case-control sampling using sequential empirical process theory. We show that the sequential empirical ROC, PPV and NPV curves converge to the sum of independent Kiefer processes and show how these results can be used to derive asymptotic results for summaries of the sequential empirical ROC, PPV and NPV curves.

Keywords and phrases

Group Sequential Methods; Empirical Process Theory; Diagnostic Testing

1. Introduction

Several recent papers have discussed the application of group sequential methodology to diagnostic biomarker studies (Tang, Emerson and Zhou, 2008; Tang and Liu, 2010; Pepe *et al.*, 2009). Group sequential study designs (i.e. study designs with multiple interim analyses) provide an opportunity to improve the efficiency of diagnostic biomarker studies by allowing studies to terminate early when the candidate marker is clearly superior or inferior to established markers or historical levels of marker performance. Many group sequential methods assume the existence of a test statistic with an independent increments covariance structure (Jennison and Turnbull, 2000). A thorough understanding of the asymptotic properties of the sequential empirical ROC, PPV and NPV curves and, specifically, verifying that their summary measures have an independent increments covariance structure, would provide great flexibility when designing group sequential diagnostic biomarker studies.

Diagnostic biomarkers are used to classify a patient as a case or a control. A dichotomous biomarker results in either a positive test, indicating that the subject should be classified as a

*NIH grant P01-CA53996

†NIH grant U01-CA86368

AMS 2000 subject classifications: Primary 62L12; secondary 62G05

case, or a negative test, indicating that the subject should be classified as a control. Many biomarkers are measured on a continuous scale and a threshold must be defined in order to translate a continuous biomarker into a positive or negative test result. Let D be a Bernoulli random variable indicating disease status with prevalence π and let X be a biomarker value with conditional distribution $F_{X|D=1}(x) = F_D(x)$ and $F_{X|D=0}(x) = F_{D^c}(x)$, where $F_D(x)$ is the distribution function for the cases and $F_{D^c}(x)$ is the distribution function for the controls. Furthermore, we define $F(x) = \pi F_D(x) + (1 - \pi) F_{D^c}(x)$ to be the biomarker distribution function for the entire population. Without loss of generality, assume that larger biomarker values are more indicative of disease. For a threshold c , a biomarker value X is translated into a positive test result if it is greater than c and a negative test result if it is less than or equal to c .

The receiver operating characteristic (ROC) curve summarizes the classification accuracy of a continuous diagnostic biomarker (Pepe, 2003) by reporting the true positive fraction (TPF) and the false positive fraction (FPF) for all possible cut-offs of the marker. For a threshold c , $TPF(c) = P[X > c | D = 1]$ and $FPR(c) = P[X > c | D = 0]$. The ROC curve is defined as,

$$ROC(c) = \{(TPE(c), FPE(c)), c \in (-\infty, \infty)\},$$

and can alternately be expressed as,

$$ROC(t) = S_D(S_D^{-1}(t)), t \in (0, 1), \quad (1.1)$$

where $S_D(x) = 1 - F_D(x)$ and $S_{D^c}(x) = 1 - F_{D^c}(x)$. $ROC(t)$ can be interpreted as the TPF corresponding to a FPF of t . Alternately, one might be interested in the inverse of the ROC curve,

$$ROC^{-1}(v) = S_{D^c}(S_D^{-1}(v)), v \in (0, 1). \quad (1.2)$$

$ROC^{-1}(\cdot)$ is indexed by the TPF and can be interpreted as the FPF corresponding to a TPF of \cdot .

The predictive accuracy of a dichotomous biomarker can be summarized by the positive predictive value (PPV) and negative predictive value (NPV).

The PPV and NPV curves were proposed as an extension of PPV and NPV to continuous markers (Moskowitz and Pepe, 2004; Zheng *et al.*, 2008). For a threshold c , $PPV(c) = P[D = 1 | X > c]$ and $NPV(c) = P[D = 0 | X \leq c]$. The PPV and NPV curves are defined as $PPV(c)$ and $NPV(c)$ for all $c \in (-\infty, \infty)$. In practice, PPV and NPV curves are indexed by a summary of the marker distribution rather than a generic threshold (Moskowitz and Pepe, 2004; Zheng *et al.*, 2008). In this paper, we consider the PPV and NPV curves indexed by the FPF and the percentile value in the entire population.

The ROC, PPV and NPV curves are commonly estimated nonparametrically to avoid making assumptions about the form of $F_D(x)$ and $F_{D^c}(x)$. This is particularly important in the case of the ROC, PPV and NPV curves because we are often interested in regions of the curve that correspond to the tails of these distributions. For example, a biomarker must possess a high specificity in order to be clinically useful in a low disease risk population screening setting, which corresponds to the upper tail of the biomarker distribution among controls.

Our understanding of the empirical ROC curve is enhanced by knowledge of its asymptotic properties. Hsieh and Turnbull (1996) showed that the empirical ROC curve converges to the sum of two independent Brownian bridges. The asymptotic normality of summary measures of the empirical ROC curve, such as the area under the ROC curve or a point on the ROC curve, can be derived from their work. To our knowledge, no asymptotic theory is available for the empirical PPV and NPV curves.

Tang, Emerson and Zhou (2008) showed that a family of weighted area under the ROC curve (wAUC) statistics has an independent increments covariance structure. It would be beneficial to show that this assumption holds for a larger class of summaries of the ROC curve. In this paper, we develop asymptotic theory for the sequential empirical ROC, PPV and NPV curves. Our results allow us to develop distribution theory for other summaries of the ROC curve and to develop distribution theory for summaries of the PPV and NPV curves.

2. Notation and Definitions

Before beginning our discussion of the sequential empirical ROC, PPV and NPV curves, we provide definitions of the sequential empirical estimates for the underlying distribution and quantile functions. Let $X_{D,1}, X_{D,2}, \dots, X_{D,n_D}$ be *i.i.d.* marker values for the cases with distribution function, $F_D(x)$, and $X_{\bar{D},1}, X_{\bar{D},2}, \dots, X_{\bar{D},n_{\bar{D}}}$ be *i.i.d.* marker values for the controls with distribution function, $F_{\bar{D}}(x)$. Furthermore, let r_D and $r_{\bar{D}}$ refer to the proportion of case and controls, respectively, that are observed at a given time point. The sequential empirical estimate of $F_D(x)$ is defined as

$$\widehat{F}_{D,r_D}(x) = \begin{cases} 0, & 0 \leq r_D < \frac{1}{n_D}, \\ \frac{1}{\lfloor r_D n_D \rfloor} \sum_{i=1}^{\lfloor r_D n_D \rfloor} 1\{X_{D,i} \leq x\}, & -\infty < x < \infty, \frac{1}{n_D} \leq r_D \leq 1, \end{cases}$$

and the sequential empirical estimate of $F_D^{-1}(t)$ is defined as

$$\widehat{F}_{D,r_D}^{-1}(t) = \begin{cases} X_{D,1, \lfloor r_D n_D \rfloor} & \text{if } t=0, 0 \leq r_D \leq 1, \\ X_{D,k, \lfloor r_D n_D \rfloor} & \text{if } \frac{k-1}{\lfloor r_D n_D \rfloor} < t \leq \frac{k}{\lfloor r_D n_D \rfloor}, 1 \leq k \leq \lfloor r_D n_D \rfloor, 0 \leq t \leq 1 \end{cases}$$

where $X_{D,1, \lfloor r_D n_D \rfloor}, X_{D,2, \lfloor r_D n_D \rfloor}, \dots, X_{D, \lfloor r_D n_D \rfloor, \lfloor r_D n_D \rfloor}$ are the sequential order statistics of the biomarker values for the cases. The sequential empirical estimates of $S_D(x)$ and $S_D^{-1}(t)$ are defined as $\widehat{S}_{D,r_D}(x) = 1 - \widehat{F}_{D,r_D}(x)$ and $\widehat{S}_{D,r_D}^{-1}(t) = \widehat{F}_{D,r_D}^{-1}(1 - t)$. The sequential empirical estimates for the control population are defined in an analogous fashion. The sequential empirical estimates of $F_{\bar{D}}(x)$ and $F_{\bar{D}}^{-1}(t)$ lead to a natural definition of the sequential empirical estimates of $F(x)$ and $F^{-1}(t)$,

$$\widehat{F}_{r_D, r_{\bar{D}}}(x) = \rho \widehat{F}_{D,r_D}(x) + (1 - \rho) \widehat{F}_{\bar{D},r_{\bar{D}}}(x)$$

and

$$\widehat{F}_{r_D, r_{\overline{D}}}^{-1}(t) = \inf \left\{ x : \widehat{F}_{r_D, r_{\overline{D}}}(x) \geq t \right\},$$

where $\widehat{F}_{r_D, r_{\overline{D}}}$ is assumed to be known. $F_{r_D, r_{\overline{D}}}(x)$ is a linear combination of $F_{D, r_D}(x)$ and $F_{\overline{D}, r_{\overline{D}}}(x)$ and is therefore indexed by both r_D , the proportion of cases observed at a given time point, and $r_{\overline{D}}$, the proportion of controls observed at a given time point.

Throughout this paper We let $0 < a < b < 1$, $0 < c < 1$, $0 < d < 1$ and make the following assumptions:

A1 $F_D(x)$ and $F_{\overline{D}}(x)$ are continuous distribution functions with continuous densities $f_D(x)$ and $f_{\overline{D}}(x)$, respectively,

A2 $f_D(x) > 0$ for $x \in (\sup\{x: F_D(x) = 0\}, \inf\{x: F_D(x) = 1\})$,

A3 $f_{\overline{D}}(x) > 0$ for $x \in (\sup\{x: F_{\overline{D}}(x) = 0\}, \inf\{x: F_{\overline{D}}(x) = 1\})$,

A4 $\frac{n_D}{n_{\overline{D}}} \rightarrow \lambda > 0$ as $n_D \rightarrow \infty$ and $n_{\overline{D}} \rightarrow \infty$, i.e. the ratio of cases to controls converges to a constant that is greater than 0.

The asymptotic results in Section 3 make use of the Kiefer process. The Kiefer process, $K(t, r)$, is a two-dimensional, mean-zero Gaussian process with covariance:

$$Cov(K(t_1, r_1), K(t_2, r_2)) = (t_1 \wedge t_2 - t_1 t_2)(r_1 \wedge r_2)$$

where \wedge represents the minimum. The Kiefer process behaves like a Brownian Bridge in t and Brownian Motion in r .

The remainder of this paper proceeds as follows. In Section 3, we develop asymptotic theory for the sequential empirical ROC, PPV and NPV curves. First, we generalize the work of Hsieh and Turnbull (1996) to the sequential empirical ROC curve by showing that the sequential empirical ROC curve converges to the sum of independent Kiefer processes. Next, we develop asymptotic theory for the sequential empirical PPV and NPV curves indexed by the FPF by writing them as functions of the sequential empirical ROC curve. Finally, we follow the approach of Pyke and Shorack (1968) to develop asymptotic theory for the PPV and NPV curves indexed by the percentile value of the marker distribution. We validate our asymptotic results by simulation in Section 4 and illustrate how they can be used to design group sequential diagnostic biomarker studies in Section 5. We conclude with a discussion in Section 6.

3. Asymptotic Results

3.1. The Sequential Empirical ROC Curve

In this section we provide asymptotic results for the sequential empirical ROC curve. Results for the inverse of the sequential empirical ROC curve are nearly identical; we direct the reader to an associated technical report for details (Koopmeiners and Feng, 2010). The sequential empirical ROC curve, $\widehat{ROC}_{r_D, r_{\overline{D}}}(t)$ is defined by substituting the sequential empirical estimates of $S_D(x)$ and $S_{\overline{D}}(x)$ into (1.1), yielding,

$$\widehat{ROC}_{r_D, r_{\bar{D}}}(t) = \widehat{S}_{D, r_D} \left(\widehat{S}_{\bar{D}, r_{\bar{D}}}^{-1}(t) \right),$$

and for ease of notation, we define,

$$R_{r_D, r_{\bar{D}}}(t) \equiv n_D^{-1/2} [n_D r_D] (\widehat{ROC}_{r_D, r_{\bar{D}}}(t) - ROC(t)).$$

The primary result in this section provides asymptotic theory for $R_{r_D, r_{\bar{D}}}(t)$. By developing asymptotic theory for $R_{r_D, r_{\bar{D}}}(t)$, we are also able to develop asymptotic theory for functionals of $R_{r_D, r_{\bar{D}}}(t)$ as a special case. Theorem 3.1 establishes the convergence of $R_{r_D, r_{\bar{D}}}(t)$ to the sum of independent Kiefer processes.

Theorem 3.1—Assume A1-A4 hold and let $\frac{f_D(S_D^{-1}(t))}{f_{\bar{D}}(S_{\bar{D}}^{-1}(t))}$ be bounded on $[a, b]$. As n_D and $n_{\bar{D}}$

$$R_{r_D, r_{\bar{D}}}(t) \rightarrow_d K_1(ROC(t), r_D) + \lambda^{1/2} \frac{r_D}{r_{\bar{D}}} \left(\frac{f_D(S_D^{-1}(t))}{f_{\bar{D}}(S_{\bar{D}}^{-1}(t))} \right) K_2(t, r_{\bar{D}})$$

uniformly for $t \in [a, b]$, $r_D \in [c, 1]$ and $r_{\bar{D}} \in [d, 1]$ where K_1 and K_2 are independent Kiefer Processes.

A proof of Theorem 3.1 can be found in the Appendix. Theorem 3.1 generalizes the results of Hsieh and Turnbull (1996) to the sequential empirical ROC curve. The proof of Theorem 3.1 is similar to the proof found in Hsieh and Turnbull (1996) but our proof relies on the more powerful sequential empirical process theory. Sequential empirical process theory generalizes asymptotic theory for the standard empirical process by introducing a parameter for time. In doing so, asymptotic results for the sequential empirical process involve the Kiefer Process. Using properties of the Kiefer Process, we are able to easily derive asymptotic results for summaries of the sequential empirical ROC curve and verify that the independent increments assumption holds in many cases. Furthermore, we can recover Hsieh and Turnbull's result as a special case of Theorem 3.1 by letting r_D and $r_{\bar{D}}$ both equal 1.

Corollary 3.2—Assume A1-A4 hold and let $\frac{f_D(S_D^{-1}(t))}{f_{\bar{D}}(S_{\bar{D}}^{-1}(t))}$ be bounded on $[a, b]$. As n_D and $n_{\bar{D}}$

$$R_{1,1}(t) \rightarrow_d B_1(ROC(t)) + \lambda^{1/2} \left(\frac{f_D(S_D^{-1}(t))}{f_{\bar{D}}(S_{\bar{D}}^{-1}(t))} \right) B_2(t)$$

uniformly for $t \in [a, b]$ where B_1 and B_2 are independent Brownian Bridges.

Proof. Immediate from Theorem 3.1 and by noting that $K(t, 1) = {}_d B(t)$.

An advantage to studying the asymptotic behavior of the sequential empirical ROC curve at the process level, rather than a single point on the sequential empirical ROC curve, is that we are able to study the joint behavior of multiple points on the ROC curve. Corollary 3.3 provides a normal approximation for a vector of points on the sequential empirical ROC curve.

Corollary 3.3—Assume A1-A4 hold and let $\frac{f_D(S_D^{-1}(t))}{f_D(S_D^{-1}(t))}$ be bounded on $[a, b]$. For t_1, t_2, \dots, t_J $(0,1)$, $r_{D,1}, r_{D,2}, \dots, r_{D,J}$ $(0,1)$ and $r_{\bar{D},1}, r_{\bar{D},2}, \dots, r_{\bar{D},J}$ $(0,1)$, a vector of arbitrary points on the sequential empirical ROC curve, $(\widehat{ROC}_{r_{D,1}, r_{\bar{D},1}}(t_1), \widehat{ROC}_{r_{D,2}, r_{\bar{D},2}}(t_2), \dots, \widehat{ROC}_{r_{D,J}, r_{\bar{D},J}}(t_J))$, is approximately multivariate normal with,

$$\widehat{ROC}_{r_{D,j}, r_{\bar{D},j}}(t_j) \sim N\left(ROC(t_j), \sigma_{\widehat{ROC}_{r_{D,j}, r_{\bar{D},j}}(t_j)}^2\right) \quad j=1, 2, \dots, J$$

where

$$\sigma_{\widehat{ROC}_{r_{D,j}, r_{\bar{D},j}}(t_j)}^2 = \frac{ROC(t_j)(1 - ROC(t_j))}{n_D r_{D,j}} + \left(\frac{f_D(S_D^{-1}(t_j))}{f_{\bar{D}}(S_{\bar{D}}^{-1}(t_j))}\right)^2 \frac{t_j(1 - t_j)}{n_{\bar{D}} r_{\bar{D},j}},$$

and

$$\begin{aligned} Cov\left[\widehat{ROC}_{r_{D,i}, r_{\bar{D},i}}(t_i), \widehat{ROC}_{r_{D,j}, r_{\bar{D},j}}(t_j)\right] &= \frac{(r_{D,i} \wedge r_{D,j})(ROC(t_i) \wedge ROC(t_j) - ROC(t_i) ROC(t_j))}{n_D r_{D,i} r_{D,j}} \\ &+ \left(\frac{f_D(S_D^{-1}(t_i))}{f_{\bar{D}}(S_{\bar{D}}^{-1}(t_i))}\right) \left(\frac{f_D(S_D^{-1}(t_j))}{f_{\bar{D}}(S_{\bar{D}}^{-1}(t_j))}\right) \frac{(r_{\bar{D},i} \wedge r_{\bar{D},j})(t_i \wedge t_j - t_i t_j)}{n_{\bar{D}} r_{\bar{D},i} r_{\bar{D},j}}. \end{aligned}$$

Proof. Immediate from Theorem 3.1.

Corollary 3.3 provides the asymptotic covariance for two points at different locations and different times on the sequential empirical ROC curve. This allows us to fully specify the joint sequential distribution of multiple points on the ROC curve, which allows us to design group sequential diagnostic biomarker studies where multiple points on the ROC curve are treated as multiple endpoints of a group sequential study. For example, we might be interested in $ROC(t_1)$ and $ROC(t_2)$, where t_1 is chosen for high specificity to rule patients in for work up and t_2 is chosen for high sensitivity to rule out patients for invasive work.

Our interest in the sequential empirical ROC curve is motivated by the need to design group sequential diagnostic biomarker studies. Our ability to design group sequential diagnostic

biomarker studies would be enhanced by showing that summaries of the sequential empirical ROC curve have an independent increments covariance structure. The simplest summary of the ROC curve is a point on the ROC curve, $ROC(t)$. $ROC(t)$ can be interpreted as the sensitivity at a specificity of $1 - t$. Corollary 3.4 shows that the sequential empirical estimator of $ROC(t)$ is asymptotically normal and has independent increments when divided its variance.

$$f_D(S_D^{-1}(t))$$

Corollary 3.4—Assume A1-A4 hold and let $f_D(S_D^{-1}(t))$ be bounded on $[a, b]$. For $t \in (0, 1)$ and J stopping times, $(\widehat{ROC}_{r_{D,1}, r_{\bar{D},1}}(t), \widehat{ROC}_{r_{D,2}, r_{\bar{D},2}}(t), \dots, \widehat{ROC}_{r_{D,J}, r_{\bar{D},J}}(t))$, is approximately multivariate normal with,

$$\widehat{ROC}_{r_{D,i}, r_{\bar{D},i}}(t) \sim N\left(ROC(t), \sigma_{\widehat{ROC}_{r_{D,i}, r_{\bar{D},i}}(t)}^2 \quad i=1, 2, \dots, J\right)$$

and

$$Cov\left[\widehat{ROC}_{r_{D,i}, r_{\bar{D},i}}(t), \widehat{ROC}_{r_{D,j}, r_{\bar{D},j}}(t)\right] = Var\left[\widehat{ROC}_{r_{D,j}, r_{\bar{D},j}}(t)\right] = \sigma_{\widehat{ROC}_{r_{D,j}, r_{\bar{D},j}}(t)}^2, r_i \leq r_j$$

where $\sigma_{\widehat{ROC}_{r_{D,j}, r_{\bar{D},j}}(t)}^2$ is defined as in Corollary 3.3.

Proof. Immediate from Corollary 3.3.

Asymptotic theory for other summary measures of the ROC curve, such as the area under the curve or the partial area under the curve, can also be derived from Theorem 3.1. This illustrates the flexibility of Theorem 3.1. By developing distribution theory for the sequential empirical ROC curve, we are able to derive distribution theory for summaries of the ROC curve as a special case.

3.2. The sequential empirical PPV and NPV curves indexed by the False Positive Fraction

In this section, we consider the sequential empirical PPV and NPV curves indexed by the false positive fraction, t . The PPV and NPV curve indexed by the false positive fraction can be written as a function of the ROC curve and their asymptotic properties can be derived using the results from Section 3.1. Asymptotic results for the PPV and NPV curve indexed by the true positive fraction, τ , can similarly be derived by writing the PPV and NPV curve as a function of the inverse of the ROC curve but are not presented in this paper. The interested reader is directed to Koopmeiners and Feng (2010) for details.

The PPV and NPV curves indexed by the false positive fraction are defined as $PPV(t) = P[D=1|X > S_D^{-1}(t)]$ and $NPV(t) = P[D=0|X \leq S_D^{-1}(t)]$ for all $t \in (0, 1)$ and can be written as functions of the ROC curve as follows:

$$PPV(t) = \frac{ROC(t)\rho}{ROC(t)\rho + t(1-\rho)}, \quad (3.1)$$

and

$$NPV(t) = \frac{(1-t)(1-\rho)}{(1-ROC(t))\rho + (1-t)(1-\rho)}. \quad (3.2)$$

The sequential empirical estimators of $PPV(t)$ and $NPV(t)$ are defined by plugging the sequential empirical estimator of $ROC(t)$ into (3.1) and (3.2), yielding,

$$\widehat{PPV}_{r_D, r_{\bar{D}}}(t) = \frac{\widehat{ROC}_{r_D, r_{\bar{D}}}(t)\rho}{\widehat{ROC}_{r_D, r_{\bar{D}}}(t)\rho + t(1-\rho)},$$

and

$$\widehat{NPV}_{r_D, r_{\bar{D}}}(t) = \frac{(1-t)(1-\rho)}{(1-\widehat{ROC}_{r_D, r_{\bar{D}}}(t))\rho + (1-t)(1-\rho)}.$$

From this point forward we only consider $\widehat{PPV}_{r_D, r_{\bar{D}}}(t)$ and note that results for $\widehat{NPV}_{r_D, r_{\bar{D}}}(t)$ are nearly identical. Again, for ease of notation, we define,

$$P_{r_D, r_{\bar{D}}}(t) \equiv n_D^{-1/2} [n_D r_D] \left(\widehat{PPV}_{r_D, r_{\bar{D}}}(t) - PPV(t) \right).$$

We begin by using the results of Section 3.1 to derive asymptotic theory for $P_{r_D, r_{\bar{D}}}(t)$. Theorem 3.5 establishes the convergence of $P_{r_D, r_{\bar{D}}}(t)$ to the sum of two independent Kiefer Processes.

Theorem 3.5—Assume A1-A4 hold and let $\frac{f_D(S_D^{-1}(t))}{f_{\bar{D}}(S_{\bar{D}}^{-1}(t))}$ be bounded on $[a, b]$. As n_D and $n_{\bar{D}}$

$$P_{r_D, r_{\bar{D}}}(t) \rightarrow d \left(\frac{t(1-\rho)\rho}{(ROC(t)\rho + t(1-\rho))^2} \right) \left(K_1(ROC(t), r_D) + \lambda^{1/2} \frac{r_D}{r_{\bar{D}}} \left(\frac{f_D(S_D^{-1}(t))}{f_{\bar{D}}(S_{\bar{D}}^{-1}(t))} \right) K_2(t, r_{\bar{D}}) \right)$$

uniformly for $t \in [a, b]$, $r_D \in [c, 1]$ and $r_{\bar{D}} \in [d, 1]$ where K_1 and K_2 are independent Kiefer Processes.

The proof of Theorem 3.5 relies on writing $P_{r_D, r_{\bar{D}}}(t)$ as a function of $R_{r_D, r_{\bar{D}}}(t)$

$$P_{r_D, r_{\bar{D}}}(t) = \frac{\left(\frac{\widehat{ROC}_{r_D, r_{\bar{D}}}(t)\rho}{\widehat{ROC}_{r_D, r_{\bar{D}}}(t)\rho + t(1-\rho)} - \frac{ROC(t)\rho}{ROC(t)\rho + t(1-\rho)} \right)}{\widehat{ROC}_{r_D, r_{\bar{D}}}(t) - ROC(t)} R_{r_D, r_{\bar{D}}}(t),$$

and applying the results of Theorem 3.1. The first term converges to $\left(\frac{t(1-\rho)\rho}{(ROC(t)\rho + t(1-\rho))^2} \right)$ and $R_{r_D, r_{\bar{D}}}(t)$ converges to the sum of two independent Keifer process by Theorem 3.1. A formal proof of Theorem 3.5 can be found in Koopmeiners and Feng (2010).

From Theorem 3.5, we can prove analogous results to Corollaries 3.3 and 3.4 for the sequential empirical PPV curve indexed by the FPF. Namely, that an arbitrary vector of points on the sequential empirical PPV curve follows a multivariate normal distribution and the sequential empirical estimate of a point on the PPV curve is approximately normally distributed with an independent increments covariance structure. We leave the formal statement of these corollaries for the Appendix but present the form of the covariance between two arbitrary points on the sequential empirical PPV curve:

$$COV \left[\widehat{PPV}_{r_{D_i^*}, r_{\bar{D}_i^*}}(t_i), \widehat{PPV}_{r_{D_j^*}, r_{\bar{D}_j^*}}(t_j) \right] = \left(\frac{t_i(1-\rho)\rho}{(ROC(t_i)\rho + t_i(1-\rho))^2} \right) \left(\frac{t_j(1-\rho)\rho}{(ROC(t_j)\rho + t_j(1-\rho))^2} \right) COV \left[\widehat{ROC}_{r_{D_i^*}, r_{\bar{D}_i^*}}(t_i), \widehat{ROC}_{r_{D_j^*}, r_{\bar{D}_j^*}}(t_j) \right].$$

$PPV(t)$ is a function of $ROC(t)$ and, therefore, distribution theory for a vector of points on the PPV curve can also be derived using the delta method and Corollary 3.3.

Asymptotic theory for the fixed-sample empirical PPV curve indexed by the FPF, which was previously unavailable, can be derived as a special case of Theorem 3.5 by letting r_D and $r_{\bar{D}}$ equal 1. The fixed-sample empirical PPV curve converges to the sum of independent Brownian bridges,

$$P_{1,1}(t) \rightarrow_d \left(\frac{t(1-\rho)\rho}{(ROC(t)\rho + t - \rho)^2} \right) \left(B_1(ROC(t)) + \lambda^{1/2} \left(\frac{f_D(S_D^{-1}(t))}{f_{\bar{D}}(S_{\bar{D}}^{-1}(t))} \right) B_2(t) \right),$$

which allows us to derive a normal approximation for the empirical estimate of a point on the PPV curve,

$$\widehat{PPV}_{1,1}(t) \sim N \left(PPV(t), \left(\frac{t(1-\rho)\rho}{(ROC(t)\rho + t - \rho)^2} \right)^2 \sigma_{\widehat{ROC}_{1,1}(t)}^2 \right)$$

where $\sigma_{\widehat{ROC}_{1,1}(t)}^2$ is defined as in Corollary 3.3.

3.3. The sequential empirical PPV and NPV curves indexed by the Percentile Value

Finally, we consider the PPV and NPV curves indexed by the proportion of the population that are classified as negative, u , and positive, $1 - u$. In this case, the PPV and NPV curves

are defined as $PPV(u) = P[D = 1 | X > F^{-1}(u)]$ and $NPV(u) = P[D = 0 | X \leq F^{-1}(u)]$ for all $u \in (0, 1)$. Under this indexing, the PPV curve can be written as

$$PPV(u) = \frac{S_D(F^{-1}(u))\rho}{1 - u}, \quad (3.3)$$

and since the NPV curve can be written as,

$$NPV(u) = \frac{u - \rho}{u} + \frac{1 - u}{u} PPV(u), \quad (3.4)$$

it suffices to study the PPV curve when considering estimation of the NPV curve.

The sequential empirical estimator of $PPV(u)$ is found by substituting the sequential empirical estimators of $S_D(x)$ and $F(x)$, along with the known value of ρ , into (3.3),

$$\widehat{PPV}_{r_D, r_{\bar{D}}}(u) = \frac{\widehat{S}_{D, r_D}(F_{r_D, r_{\bar{D}}}^{-1}(u))\rho}{1 - u}, \quad (3.5)$$

and the sequential empirical estimator of $NPV(u)$ is found by substituting the sequential empirical estimator of $PPV(u)$ into (3.4),

$$\widehat{NPV}_{r_D, r_{\bar{D}}}(u) = \frac{u - \rho}{u} + \frac{1 - u}{u} \widehat{PPV}_{r_D, r_{\bar{D}}}(u). \quad (3.6)$$

Finally, we define,

$$P_{r_D, r_{\bar{D}}}(u) = n_D^{-1/2} [n_D r_D] \left(\widehat{PPV}_{r_D, r_{\bar{D}}}(u) - PPV(u) \right)$$

and

$$N_{r_D, r_{\bar{D}}}(u) = n_D^{-1/2} [n_D r_D] \left(\widehat{NPV}_{r_D, r_{\bar{D}}}(u) - NPV(u) \right)$$

for mathematical convenience. We begin by developing distribution theory for $P_{r_D, r_{\bar{D}}}(u)$. Theorem 3.6 establishes the convergence of the sequential empirical PPV curve to the sum of two independent Kiefer Processes.

Theorem 3.6—Assume A1 - A4 hold and let $f_D(F^{-1}(u)) f(F^{-1}(u))$ be bounded on $[a, b]$. As $n_D \rightarrow \infty$ and $n_{\bar{D}} \rightarrow \infty$

$$P_{r_D, r_{\bar{D}}}(u) \rightarrow_d -\frac{\rho(1 - \rho) f_{\bar{D}}(F^{-1}(u))}{1 - u} K_1(F_D(F^{-1}(u)), r_D) + \frac{\rho(1 - \rho) f_D(F^{-1}(u))}{1 - u} \frac{f(F^{-1}(u))}{f(F^{-1}(u))} \sqrt{\lambda} \frac{r_D}{r_{\bar{D}}} K_2(F_{\bar{D}}(F^{-1}(u)), r_{\bar{D}})$$

uniformly for $u \in [a, b]$, $r_D \in [c, 1]$ and $r_{\bar{D}} \in [d, 1]$ where K_1 and K_2 are independent Kiefer Processes.

The proof of Theorem 3.6 is complicated by the fact that $S_{D,r_D}(x)$ and $\widehat{F}_{r_D,r_{\bar{D}}}^{-1}(t)$ are correlated because $\widehat{F}_{D,r_D}(x)$ is a linear combination of $F_{D,r_D}(x)$ and $F_{\bar{D},r_{\bar{D}}}(x)$. In contrast, the sequential empirical ROC curve and the sequential empirical PPV curve indexed by the FPF are functionals of two independent sequential empirical estimators, $S_{D,r_D}(x)$ and $\widehat{S}_{\bar{D},r_{\bar{D}}}^{-1}(t)$, which makes it easier to show that $R_{r_D,r_D}(t)$ and $P_{r_D,r_D}(t)$ converge to the sum of independent Kiefer processes. To account for the correlation between $S_{D,r_D}(x)$ and $\widehat{F}_{r_D,r_{\bar{D}}}^{-1}(t)$ we follow the approach of Pyke and Shorack (1968), who prove a similar result for two correlated, fixed-sample empirical processes. The proof of Theorem 3.6 can be found in the Appendix.

Theorem 3.6 also establishes asymptotic theory for the sequential empirical NPV curve because $\widehat{NPV}_{r_D,r_{\bar{D}}}(t)$ is a function of $\widehat{PPV}_{r_D,r_{\bar{D}}}(t)$. Corollary 3.7 establishes the convergence of $N_{r_D,r_{\bar{D}}}(t)$ to the sum of two independent Kiefer processes.

Corollary 3.7—Assume A1 - A4 hold and let $\frac{f_D(F^{-1}(u))}{f(F^{-1}(u))}$ be bounded on $[a, b]$. As n_D and $n_{\bar{D}}$

$$N_{r_D,r_{\bar{D}}}(u) \rightarrow_d -\frac{\rho(1-\rho)}{1-u} \frac{f_D(F^{-1}(u))}{f(F^{-1}(u))} K_1(F_D(F^{-1}(u)), r_D) + \frac{\rho(1-\rho)}{1-u} \frac{f_D(F^{-1}(u))}{f(F^{-1}(u))} \sqrt{\lambda \frac{r_D}{r_{\bar{D}}}} K_2(F_{\bar{D}}(F^{-1}(u)), r_{\bar{D}})$$

uniformly for $u \in [a, b]$, $r_D \in [c, 1]$ and $r_{\bar{D}} \in [d, 1]$ where K_1 and K_2 are independent Kiefer Processes.

Corollary 3.7 is immediate from Theorem 3.6 by noting that,

$$N_{r_D,r_{\bar{D}}}(t) = \frac{1-u}{u} P_{r_D,r_{\bar{D}}}(t).$$

As with the ROC curve and the PPV curve indexed by the FPF, Theorem 3.6 and Corollary 3.7 allow us to develop distribution theory for summaries of the PPV and NPV curve indexed by u . Distribution theory for a vector of points on the PPV or NPV curve is left for the Appendix but we choose to highlight the joint distribution of the sequential empirical estimate of a single point on the PPV or NPV curve. Corollary 3.8 establishes that the sequential empirical estimate of a point on the PPV or NPV curve is asymptotically normal and has independent increments when divided by its variance.

Corollary 3.8—Assume A1 - A4 hold and let $\frac{f_D(F^{-1}(u))}{f(F^{-1}(u))}$ be bounded on $[a, b]$. For $u \in (0, 1)$ and J stopping times,

A. $(\widehat{PPV}_{r_{D,1},r_{\overline{D},1}}(u), \widehat{PPV}_{r_{D,2},r_{\overline{D},2}}(u), \dots, \widehat{PPV}_{r_{D,J},r_{\overline{D},J}}(u))$, is approximately multivariate normal with,

$$\widehat{PPV}_{r_{D,i},r_{\overline{D},i}}(u) \sim N\left(PPV(u), \sigma_{\widehat{PPV}_{r_{D,i},r_{\overline{D},i}}(u)}^2\right) i=1, 2, \dots, J$$

and

$$COV\left[\widehat{PPV}_{r_{D,i},r_{\overline{D},i}}(u), \widehat{PPV}_{r_{D,j},r_{\overline{D},j}}(u)\right] = V ar\left[\widehat{PPV}_{r_{D,j},r_{\overline{D},j}}(u)\right] = \sigma_{\widehat{PPV}_{r_{D,j},r_{\overline{D},j}}(u)}^2, r_i \leq r_j$$

where

$$\sigma_{\widehat{PPV}_{r_{D,j},r_{\overline{D},j}}(u)}^2 = \frac{\left(\frac{f_D(F^{-1}(u))}{f(F^{-1}(u))}(1-\rho)\right)^2 PPV(u) \left(\frac{\rho}{1-u} - PPV(u)\right) + \left(\frac{f_D(F^{-1}(u))}{f(F^{-1}(u))}\rho\right)^2 (1-PPV(u)) \left(\frac{u-\rho}{1-u} + PPV(u)\right)}{n_D r_{D,j} + n_{\overline{D}} r_{\overline{D},j}}$$

B. $(\widehat{NPV}_{r_{D,1},r_{\overline{D},1}}(u), \widehat{NPV}_{r_{D,2},r_{\overline{D},2}}(u), \dots, \widehat{NPV}_{r_{D,J},r_{\overline{D},J}}(u))$, is approximately multivariate normal with,

$$\widehat{NPV}_{r_{D,i},r_{\overline{D},i}}(u) \sim N\left(NPV(u), \sigma_{\widehat{NPV}_{r_{D,i},r_{\overline{D},i}}(u)}^2\right) i=1, 2, \dots, J$$

and

$$COV\left[\widehat{NPV}_{r_{D,i},r_{\overline{D},i}}(u), \widehat{NPV}_{r_{D,j},r_{\overline{D},j}}(u)\right] = V ar\left[\widehat{NPV}_{r_{D,j},r_{\overline{D},j}}(u)\right] = \sigma_{\widehat{NPV}_{r_{D,j},r_{\overline{D},j}}(u)}^2, r_i \leq r_j$$

where

$$\sigma_{\widehat{NPV}_{r_{D,j},r_{\overline{D},j}}(u)}^2 = \frac{\left(\frac{f_D(F^{-1}(u))}{f(F^{-1}(u))}(1-\rho)\right)^2 (NPV(u) + \frac{\rho-u}{u})(1-NPV(u)) + \left(\frac{f_D(F^{-1}(u))}{f(F^{-1}(u))}\rho\right)^2 (1-NPV(u)) \left(\frac{1-\rho}{u} + NPV(u)\right)}{n_D r_{D,j} + n_{\overline{D}} r_{\overline{D},j}}$$

It is immediate from Theorem 3.6 and Corollary 3.7 that $\widehat{PPV}_{r_{D},r_{\overline{D}}}(u)$ and $\widehat{NPV}_{r_{D},r_{\overline{D}}}(u)$ are asymptotically normal with an independent increments covariance structure. By noting that,

$$F_D(F^{-1}(u)) = 1 - \frac{1-u}{\rho} PPV(u) = \frac{u}{\rho} (1 - NPV(u)),$$

and

$$F_{\bar{D}}(F^{-1}(u)) = 1 - \frac{1-u}{1-\rho} (1 - PPV(u)) = \frac{u}{1-\rho} NPV(u),$$

we can write the asymptotic variances of $\widehat{PPV}_{r_D, r_{\bar{D}}}(u)$ and $\widehat{NPV}_{r_D, r_{\bar{D}}}(u)$ as functions of $PPV(u)$ and $NPV(u)$, respectively. This provides a better understanding of the mean-variance relationship for the asymptotic distributions of $\widehat{PPV}_{r_D, r_{\bar{D}}}(u)$ and $\widehat{NPV}_{r_D, r_{\bar{D}}}(u)$ and, perhaps, provides a form of the variance that is easier to work with in practical situations (i.e. study design, estimating the standard error, etc.).

An important component of Theorem 3.6 and Corollary 3.7 is that not only do $Pr_{D, r_D}(u)$ and $Nr_{D, r_D}(u)$ converge to the sum of independent Kiefer processes, but they both converge to the same two Kiefer processes. As a result, we are able to derive the correlation between a point on the PPV curve and a point on the NPV curve. Corollary 3.9 provides a bivariate normal approximation for a point on the PPV and a point on the NPV curve.

Corollary 3.9—Assume A1 - A4 hold and let $\frac{f_D(F^{-1}(u))}{f(F^{-1}(u))}$ be bounded on $[a, b]$. For $u_1, u_2 \in (0, 1)$, $(\widehat{PPV}_{r_{D,1}, r_{\bar{D},1}}(u_1), \widehat{NPV}_{r_{D,2}, r_{\bar{D},2}}(u_2))$, is approximately bivariate normally distributed with,

$$\widehat{PPV}_{r_{D,1}, r_{\bar{D},1}}(u_1) \sim N\left(PPV(u_1), \sigma_{\widehat{PPV}_{r_{D,1}, r_{\bar{D},1}}(u_1)}^2\right),$$

and

$$\widehat{NPV}_{r_{D,2}, r_{\bar{D},2}}(u_2) \sim N\left(NPV(u_2), \sigma_{\widehat{NPV}_{r_{D,2}, r_{\bar{D},2}}(u_2)}^2\right),$$

with,

$$\begin{aligned}
 COV & \left[\widehat{PPV}_{r_{D,1}, r_{\bar{D},1}}(u_1), \widehat{NPV}_{r_{D,2}, r_{\bar{D},2}}(u_2) \right] \\
 &= \frac{(1-\rho)^2 u_1 (1-u_2) \frac{f_{\bar{D}}(F^{-1}(u_1))}{f(F^{-1}(u_1))} \frac{f_{\bar{D}}(F^{-1}(u_2))}{f(F^{-1}(u_2))} (r_{D,1} \wedge r_{D,2}) (1 - NPV(u_1)) PPV(u_2)}{(1-u_1) u_2 n_D r_{D,1} r_{D,2}} \\
 &+ \frac{\rho^2 u_1 (1-u_2) \frac{f_D(F^{-1}(u_1))}{f(F^{-1}(u_1))} \frac{f_D(F^{-1}(u_2))}{f(F^{-1}(u_2))} (r_{\bar{D},2} \wedge r_{\bar{D},2}) NPV(u_1) (1 - PPV(u_2))}{(1-u_1) u_2 n_D r_{D,1} r_{D,2}},
 \end{aligned}$$

when $u_1 \geq u_2$ and,

$$\begin{aligned}
 COV & \left[\widehat{PPV}_{r_{D,1}, r_{\bar{D},1}}(u_1), \widehat{NPV}_{r_{D,2}, r_{\bar{D},2}}(u_2) \right] \\
 &= (1-\rho)^2 \frac{f_{\bar{D}}(F^{-1}(u_1))}{f(F^{-1}(u_1))} \frac{f_{\bar{D}}(F^{-1}(u_2))}{f(F^{-1}(u_2))} (r_{D,1} \wedge r_{D,2}) \frac{(1 - NPV(u_2)) PPV(u_1)}{n_D r_{D,1} r_{D,2}} \\
 &+ \rho^2 \frac{f_D(F^{-1}(u_1))}{f(F^{-1}(u_1))} \frac{f_D(F^{-1}(u_2))}{f(F^{-1}(u_2))} (r_{\bar{D},2} \wedge r_{\bar{D},2}) \frac{NPV(u_2) (1 - PPV(u_1))}{n_D r_{D,1} r_{D,2}},
 \end{aligned}$$

when $u_2 \geq u_1$, where $\sigma_{\widehat{PPV}_{r_{D,1}, r_{\bar{D},1}}(u_1)}^2$ and $\sigma_{\widehat{NPV}_{r_{D,2}, r_{\bar{D},2}}(u_2)}^2$ are defined as in Corollary 3.8.

The case of a point on the PPV curve and a point on the NPV curve is presented for simplicity but Corollary 3.9 can be extended to an arbitrary vector of points on the PPV and NPV curves. Corollary 3.9 has obvious practical implications. It is not uncommon to classify the bottom $u_1 \times 100\%$ of the population as “low-risk”, the top $(1 - u_2) \times 100\%$ of the population as “high-risk” and the remainder of the population as “moderate-risk”. In this case, one would be interested in the NPV of the low-risk group and the PPV of the high-risk group. Corollary 3.9 provides the joint convergence of these two estimates.

Finally, we note that asymptotic results for the fixed-sample empirical PPV and NPV curves indexed by the percentile value of the marker distribution can be derived as a special case of the results in this section. It is immediate from Theorem 3.6 and Corollary 3.7 that the fixed-sample empirical PPV and NPV curves converge to the sum of independent Brownian Bridges by letting r_D and $r_{\bar{D}}$ both equal 1. Furthermore, Corollary 3.8 provides a normal approximation for the fixed-sample empirical estimate of a point on the PPV or NPV curve for the special case when $J = 1$.

4. Finite Sample Properties

A simulation study was completed to assess the finite sample properties of the results in Theorems 3.1, 3.5 and 3.6. We simulated 10,000 studies with n_D controls and n_D cases. Biomarker values for the controls were drawn from a standard normal distribution and biomarker values for the cases were drawn from a normal distribution with mean and standard deviation equal to 1. A prevalence of .2 was used for estimation of the PPV curve. Figure 1 presents the true ROC and PPV curves for this scenario. For each realization, we calculated $R_{r_D, r_{\bar{D}}}(t)$, $Pr_{r_D, r_{\bar{D}}}(t)$ and $Pr_{r_D, r_{\bar{D}}}(u)$ and evaluated the expected value, normality and covariance for various combinations of r_D , $r_{\bar{D}}$ and t or u . Normality was evaluated by providing a summary of information found in a normal q-q plot. Instead of providing the entire plot, we provide the (simulated) probability of being less than the 5th, 25th, 50th, 75th

and 95th percentile of a normal distribution with variance derived using the results in Theorems 3.1, 3.5 and 3.6. Similarly, the simulated covariance matrices were compared to the theoretical covariance matrices derived using the results in Theorems 3.1, 3.5 and 3.6.

Table 1 presents simulation results for $Rr_{D,r_D}(t)$. The expected value was close to 0 in all cases with only a small amount of bias observed when $t = .2$. The probability of being less than the theoretical 5th and 95th percentile was close to the nominal value for all sample sizes, while the probability of being less than the 25th, 50th and 75th percentile was less than the nominal value with 50 cases and 50 controls but approached the correct values as sample size increases. The observed variance and covariance were less than expected with 50 cases and 50 controls but the observed covariance matrix approached the theoretical covariance matrix in larger sample sizes. This phenomenon is likely due to the sample space for $ROC(t)$ being restricted to the unit interval. $\widehat{ROC}(t)$ is less likely to equal 0 or 1 as sample size increases and the normal approximation will be more accurate. Similar results were observed for $Pr_{D,r_D}(t)$ and $Pr_{D,r_D}(u)$ but were omitted for brevity.

5. Application

The results of Section 3 provide fundamental theory that allows existing group sequential methodology to be applied to summaries of the ROC, PPV and NPV curves. In this section, we present an example of how these results can be used to design group sequential diagnostic biomarker studies. Our application is presented in the context of a study to evaluate the diagnostic accuracy of des-gamma carboxyprothrombin (DCP), a novel biomarker for the early detection of hepatocellular carcinoma (HCC). A multi-center study was completed to compare the diagnostic accuracy of DCP to that of alpha-fetoprotein (AFP), the most widely used biomarker for the detection of HCC (Marrero *et al.*, 2009) but in our application we will only consider the design of a study to compare DCP to historical levels of diagnostic accuracy for AFP.

We consider a study to evaluate the predictive accuracy of DCP using the following novel design that makes use of the joint asymptotic theory for the PPV and NPV curve derived in Section 3.8. Assume that the prevalence of HCC in the population of interest is 0.2. In this case, one might call the bottom 60% percent of biomarker values “negative”, the top 10% of the biomarker values “positive” and refer the remaining subjects for further evaluation. Under this scenario, we would desire a high NPV for negative test results, $NPV(0.6)$, and a high PPV for positive test results, $PPV(0.9)$. The $NPV(0.6)$ for AFP is 0.92 and the $PPV(0.9)$ is 0.82. To determine if DCP improves on the predictive accuracy of AFP, we would test the hypothesis,

$$H_0: NPV(0.6) \leq 0.9 \text{ or } PPV(0.9) \leq 0.8,$$

versus,

$$H_a: NPV(0.6) > 0.9 \text{ and } PPV(0.9) > 0.8,$$

using the test statistics, $Z_{NPV(u_1)}$ and $Z_{PPV(u_2)}$, where $Z_{NPV(u_1)}$ is defined as

$$Z_{NPV(u_1)} = \frac{\widehat{NPV}(0.6) - NPV(0.6)_0}{\sigma_{NPV(0.6)_0}},$$

and $Z_{PPV(u_2)}$ is defined in an analogous fashion.

We consider a group sequential design using the error spending approach proposed by Hwang, Shih and De Cani (1990). The overall null hypothesis will only be rejected if the null hypotheses for both $NPV(0.6)$ and $PPV(0.9)$ are rejected. In the context of a group sequential study, this means that the study will stop early to reject the null hypothesis if $Z_{NPV(u_1)}$ and $Z_{PPV(u_2)}$ both cross the boundary for rejecting the null hypothesis but the study will stop early for futility if either $Z_{NPV(u_1)}$ or $Z_{PPV(u_2)}$ cross the futility boundary. This implies that we do not need to adjust the type-I error rate to account for multiple endpoints but we do need to consider the joint probability of rejecting the null hypothesis when determining the power.

The sample size for our study is chosen to achieve 90% power under the alternative hypothesis $NPV(0.6) = 0.95$ and $PPV(0.9) = 0.90$. A closed-form formula for determining the required sample size is not available. Instead, the sample size for a fixed sample design is derived by numerically solving,

$$P(Z_{NPV(u_1)} > Z_{1-\alpha/2}, PPV(u_2) > Z_{1-\alpha/2} | NPV(u_1) = 0.95, PPV(u_2) = 0.90),$$

for n_D , where the joint distribution of $Z_{NPV(u_1)}$ and $Z_{PPV(u_2)}$ is derived by applying the delta method to the joint asymptotic normal distribution of $\widehat{NPV}_{r_D, r_{\bar{D}}}(u_1)$ and $\widehat{PPV}_{r_D, r_{\bar{D}}}(u_2)$ found in Corollary 3.9. Assuming a one-to-one ratio of cases to controls, 702 cases are required to achieve 90% power under the alternative hypothesis. This sample size must be multiplied by an inflation factor to determine the maximum sample size for a group sequential design (i.e. the sample size if the study does not stop at the interim analyses) in order for the group sequential design to maintain the same type-I error rate and power as the fixed-sample design (Jennison and Turnbull, 2000). Using the `gsDesign` package in R, we find that the maximum sample size for group sequential studies with two, three and four stopping times are 724, 737 and 745 cases, respectively. However, as illustrated in the simulation which follows, the actual sample sizes required in group sequential studies are generally smaller than these maximum values.

Table 2 presents simulation results using a fixed-sample design and group sequential designs with two, three and four stopping times. Biomarker values for the controls were simulated from a standard normal distribution and biomarker values for the cases were simulated from a normal distribution with mean and variance chosen to achieve the desired value of $NPV(0.6)$ and $PPV(0.9)$. The advantages of group sequential designs are clear. The group sequential designs have similar type-I error rate and power to the fixed-sample design but with substantially smaller expected sample sizes in all scenarios.

6. Discussion

In this paper, we derived asymptotic properties of the sequential empirical ROC, PPV and NPV curves. We first extended the work of Hsieh and Turnbull (1996) to the sequential empirical ROC curve and used these results to develop distribution theory for summaries of the sequential empirical ROC curve. Next, we considered asymptotic theory for the

sequential empirical PPV curve indexed by the FPF and percentile value in the entire population. These results were used to develop distribution theory for summaries of the sequential empirical PPV curve. Asymptotic theory for the fixed-sample PPV curve, which was previously unavailable, was developed as a special case.

This work was motivated by the desire to design group sequential diagnostic biomarker studies. In Section 5, we illustrated how our results can be used to design group sequential diagnostic biomarker studies. Our simulation results clearly illustrate the advantages of group sequential designs. In both cases, the group sequential designs have similar type-I error rate and power than the fixed-sample designs but with substantially smaller expected sample size.

An advantage to our approach is that we are able to investigate the joint behavior of multiple points on the ROC and PPV curve. The primary end-point of a diagnostic biomarker study may be a single point on the ROC or PPV curve but other points on the ROC or PPV curve may also be of interest. The results of Theorems 3.1, 3.5 and 3.6 allow us to apply existing group sequential methodology for analyzing multiple endpoints to scenarios where multiple points on the ROC or PPV curve are of interest in a group sequential diagnostic biomarker study (Liu and Hall, 2001).

We considered estimation of the sequential empirical ROC and PPV curve under case-control sampling. The asymptotic properties of the sequential empirical ROC and PPV curve under other sampling schemes are also of interest. We are currently working on extending the results of this paper to estimation of the sequential empirical ROC and PPV curve under cohort and nested case-control sampling.

The theory developed in this paper applies to sequential testing of the diagnostic accuracy of a continuous test. In many cases, diagnostic tests take the form of multi-level ordinal data (cancer staging, for example). Methods exist extending the ROC curve to ordinal data (Dorfman and Alf, 1960) but further work is needed to verify that group sequential methods can be applied in these settings.

Response adaptive clinical trials have been proposed as a means to provide greater flexibility when designing therapeutic clinical trials. Response adaptive clinical trials adjust the design characteristics of the study (sample size, percent randomized to each group, etc.) in response to outcomes for subjects enrolled earlier in the study. Recently, Zhu and Hu (2010) showed that a class of test statistics from a response adaptive clinical trial converges to Brownian Motion when considered sequentially (similar to what we have shown for the empirical ROC, PPV and NPV curves), which allows existing group sequential methodology to be applied to response adaptive clinical trials. Future work will be needed to consider how response adaptive designs can be applied in the setting of group sequential diagnostic biomarker studies.

Acknowledgments

The authors would like to thank Jon Wellner for his assistance with the technical aspects of sequential empirical process theory and Julian Wolfson for his thoughtful comments on the manuscript. In addition, we would also like to thank the Associate Editor and Referee for their helpful comments that greatly improved the manuscript.

Appendix A

Supplementary Results for Section 3

A.1. Supplementary Results for Section 3.1

Proof of Theorem 3.1—First, note that

$$\begin{aligned} n_D^{-1/2} [n_D r_D] (R\widehat{OC}_{r_D, r_{\bar{D}}}(t) - ROC(t)) &= n_D^{-1/2} [n_D r_D] \left(\widehat{S}_{r_D, r_{\bar{D}}}(\widehat{S}_{\bar{D}, r_{\bar{D}}}^{-1}(t)) - S_D(\widehat{S}_{\bar{D}}^{-1}(t)) \right) \\ &= n_D^{-1/2} [n_D r_D] \left(\widehat{S}_{r_D, r_{\bar{D}}}(\widehat{S}_{\bar{D}, r_{\bar{D}}}^{-1}(t)) - S_D(\widehat{S}_{\bar{D}, r_{\bar{D}}}^{-1}(t)) \right) \\ &\quad + n_D^{-1/2} [n_D r_D] \left(S_D(\widehat{S}_{\bar{D}, r_{\bar{D}}}^{-1}(t)) - S_D(\widehat{S}_{\bar{D}}^{-1}(t)) \right). \end{aligned}$$

The first term converges to a Kiefer process. We note that

$$\begin{aligned} &\sup_{c \leq r_D \leq 1} \sup_{d \leq r_{\bar{D}} \leq 1} \sup_{a \leq t \leq b} \left| F_{\bar{D}} \left(\widehat{F}_{\bar{D}, r_{\bar{D}}}^{-1}(t) \right) \right. \\ &\quad - t \left| = \frac{n_{\bar{D}}}{[n_{\bar{D}} d]} \sup_{c \leq r_D \leq 1} \sup_{d \leq r_{\bar{D}} \leq 1} \sup_{a \leq t \leq b} \frac{[n_{\bar{D}} d]}{n_{\bar{D}}} \left| F_{\bar{D}} \left(\widehat{F}_{\bar{D}, r_{\bar{D}}}^{-1}(t) \right) \right. \right. \\ &\quad \left. \left. - t \right| \leq \frac{n_{\bar{D}}}{[n_{\bar{D}} d]} \sup_{c \leq r_D \leq 1} \sup_{d \leq r_{\bar{D}} \leq 1} \sup_{a \leq t \leq b} \frac{[n_{\bar{D}} r_{\bar{D}}]}{n_{\bar{D}}} \left| F_{\bar{D}} \left(\widehat{F}_{\bar{D}, r_{\bar{D}}}^{-1}(t) \right) - t \right|. \end{aligned}$$

Therefore

$$\sup_{c \leq r_D \leq 1} \sup_{d \leq r_{\bar{D}} \leq 1} \sup_{a \leq t \leq b} \left| F_{\bar{D}} \left(\widehat{F}_{\bar{D}, r_{\bar{D}}}^{-1}(t) \right) - t \right| \xrightarrow{a.s.} 0 \quad (\text{A.1})$$

by the Glivenko-Cantelli Theorems (1.51 and 1.52 in Csörgö and Szyszkowicz (1998)) and

because $\frac{n_{\bar{D}}}{[n_{\bar{D}} d]} \rightarrow \frac{1}{d}$. Furthermore, $F_{\bar{D}}^{-1}(t)$ will be continuous by A1-A3 and will be uniformly continuous on $[a, b]$. Therefore,

$$\sup_{c \leq r_D \leq 1} \sup_{d \leq r_{\bar{D}} \leq 1} \sup_{a \leq t \leq b} \left| \widehat{F}_{\bar{D}, r_{\bar{D}}}^{-1}(t) - F_{\bar{D}}^{-1}(t) \right| \xrightarrow{a.s.} 0. \quad (\text{A.2})$$

We note that due to the continuity of $F_D(x)$, $S_{\bar{D}}^{-1}(t) = F_{\bar{D}}^{-1}(1 - t)$ and therefore (A.2) also applies to $S_{\bar{D}}^{-1}(t)$. From corollary 1.A in Csörgö and Szyszkowicz (1998), (A.2) and the uniform continuity of the Kiefer process, we have

$$n_D^{-1/2} [n_D r_D] \left(\widehat{S}_{D, r_D}^{-1}(\widehat{S}_{D, r_D}^{-1}(t)) - S_D(\widehat{S}_{D, r_D}^{-1}(t)) \right) \rightarrow_d K_1(ROC(t), r_D). \quad (A.3)$$

The second term can be re-written as

$$\begin{aligned} & n_D^{-1/2} [n_D r_D] \left(S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) - S_D \left(S_D^{-1}(t) \right) \right) \\ &= \frac{n_D^{-1/2} [n_D r_D]}{n_D^{-1/2} [n_D r_D]} \frac{\left(S_D \left(S_D^{-1} \left(S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) \right) \right) - S_D \left(S_D^{-1}(t) \right) \right)}{S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) - t} n_D^{-1/2} [n_D r_D] \left(S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) - \widehat{S}_{D, r_D} \left(\widehat{S}_{D, r_D}^{-1}(t) \right) \right) \\ &+ \frac{n_D^{-1/2} [n_D r_D]}{n_D^{-1/2} [n_D r_D]} \frac{\left(S_D \left(S_D^{-1} \left(S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) \right) \right) - S_D \left(S_D^{-1}(t) \right) \right)}{S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) - t} n_D^{-1/2} [n_D r_D] \left(\widehat{S}_{D, r_D} \left(\widehat{S}_{D, r_D}^{-1}(t) \right) - t \right). \end{aligned}$$

By the mean value theorem, there exists a $S_D^{-1} \left(\widehat{S}_{D, r_D}^{-1}(t) \right)$ between $S_D^{-1} \left(\widehat{S}_{D, r_D}^{-1}(t) \right)$ and t such that

$$\frac{S_D \left(S_D^{-1} \left(S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) \right) \right) - S_D \left(S_D^{-1}(t) \right)}{S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) - t} = \frac{f_D \left(S_D^{-1} \left(S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) \right) \right)}{f_D \left(S_D^{-1} \left(S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) \right) \right)}.$$

From (A.1), we know that $S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) \rightarrow_{a.s.} t$, uniformly for $t \in [a, b]$, $r_D \in [c, 1]$ and r_D

$\in [d, 1]$, and, therefore, $S_D^{-1} \left(\widehat{S}_{D, r_D}^{-1}(t) \right) \rightarrow_{a.s.} t$, uniformly for $t \in [a, b]$, $r_D \in [c, 1]$ and $r_D \in [d, 1]$.

$$\frac{f_D \left(S_D^{-1}(t) \right)}{f_D \left(S_D^{-1}(t) \right)}$$

This, along with the uniform continuity of $f_D \left(S_D^{-1}(t) \right)$, allows us to conclude that

$$\sup_{c \leq r_D \leq 1} \sup_{d \leq r_D \leq 1} \sup_{a \leq t \leq b} \left| \frac{f_D \left(S_D^{-1} \left(S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) \right) \right)}{f_D \left(S_D^{-1} \left(S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) \right) \right)} - \frac{f_D \left(S_D^{-1}(t) \right)}{f_D \left(S_D^{-1}(t) \right)} \right| \rightarrow_{a.s.} 0,$$

which implies,

$$\sup_{c \leq r_D \leq 1} \sup_{d \leq r_D \leq 1} \sup_{a \leq t \leq b} \left| \frac{S_D \left(S_D^{-1} \left(S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) \right) \right) - S_D \left(S_D^{-1}(t) \right)}{S_D \left(\widehat{S}_{D, r_D}^{-1}(t) \right) - t} - \frac{f_D \left(S_D^{-1}(t) \right)}{f_D \left(S_D^{-1}(t) \right)} \right| \rightarrow_{a.s.} 0. \quad (A.4)$$

For all $r_D \in [d, 1]$,

$$\sup_{a \leq t \leq b} \left| \widehat{S}_{\bar{D}, r_{\bar{D}}} \left(\widehat{S}_{\bar{D}, r_{\bar{D}}}^{-1}(t) \right) - t \right| \leq_{a.s.} \frac{1}{\lfloor n_{\bar{D}} r_{\bar{D}} \rfloor}.$$

Therefore,

$$\sup_{c \leq r_{\bar{D}} \leq 1} \sup_{d \leq r_{\bar{D}} \leq 1} \sup_{a \leq t \leq b} n_{\bar{D}}^{-1/2} \lfloor n_{\bar{D}} r_{\bar{D}} \rfloor \left| \widehat{S}_{\bar{D}, r_{\bar{D}}} \left(\widehat{S}_{\bar{D}, r_{\bar{D}}}^{-1}(t) \right) - t \right| \leq_{a.s.} \frac{1}{n_{\bar{D}}^{1/2}},$$

and

$$\sup_{c \leq r_{\bar{D}} \leq 1} \sup_{d \leq r_{\bar{D}} \leq 1} \sup_{a \leq t \leq b} n_{\bar{D}}^{-1/2} \lfloor n_{\bar{D}} r_{\bar{D}} \rfloor \left| \widehat{S}_{\bar{D}, r_{\bar{D}}} \left(\widehat{S}_{\bar{D}, r_{\bar{D}}}^{-1}(t) \right) - t \right| \leq_{a.s.} 0. \quad (A.5)$$

From corollary 1.A in Csörgő and Szyszkowicz (1998), (A.2) and the uniform continuity of the Kiefer process, we have

$$n_{\bar{D}}^{-1/2} \lfloor n_{\bar{D}} r_{\bar{D}} \rfloor \left(S_{\bar{D}} \left(\widehat{S}_{\bar{D}, r_{\bar{D}}}^{-1}(t) \right) - \widehat{S}_{\bar{D}, r_{\bar{D}}} \left(\widehat{S}_{\bar{D}, r_{\bar{D}}}^{-1}(t) \right) \right) \rightarrow_d K_2(t, r_{\bar{D}}). \quad (A.6)$$

By (A.4), (A.5), (A.6) and noting that $\frac{n_{\bar{D}}^{-1/2} \lfloor n_{\bar{D}} r_{\bar{D}} \rfloor}{n_{\bar{D}}^{-1/2} \lfloor n_{\bar{D}} r_{\bar{D}} \rfloor} \rightarrow \lambda^{1/2} \frac{r_{\bar{D}}}{r_{\bar{D}}}$, we conclude that

$$n_{\bar{D}}^{-1/2} \lfloor n_{\bar{D}} r_{\bar{D}} \rfloor \left(S_{\bar{D}} \left(\widehat{S}_{\bar{D}, r_{\bar{D}}}^{-1}(t) \right) - S_{\bar{D}} \left(S_{\bar{D}}^{-1}(t) \right) \right) \rightarrow_d \lambda^{1/2} \frac{r_{\bar{D}}}{r_{\bar{D}}} \left(\frac{f_{\bar{D}} \left(S_{\bar{D}}^{-1}(t) \right)}{f_{\bar{D}} \left(S_{\bar{D}}^{-1}(t) \right)} \right) K_2(t, r_{\bar{D}}). \quad (A.7)$$

Summing (A.3) and (A.7) gives the desired result.

A.2. Supplementary Results for Section 3.2

Corollary A.1

$$\frac{f_{\bar{D}} \left(S_{\bar{D}}^{-1}(t) \right)}{f_{\bar{D}} \left(S_{\bar{D}}^{-1}(t) \right)}$$

Assume A1-A4 hold and let $f_{\bar{D}} \left(S_{\bar{D}}^{-1}(t) \right)$ be bounded on $[a, b]$. For $t_1, t_2, \dots, t_J \in (0, 1)$, $r_{D,1}, r_{D,2}, \dots, r_{D,J} \in (0, 1]$ and $r_{D,1}, r_{D,2}, \dots, r_{D,J} \in (0, 1]$, a vector of arbitrary points on the sequential empirical PPV curve, $\left(\widehat{PPV}_{r_{D,1}, r_{\bar{D},1}}(t_1), \widehat{PPV}_{r_{D,2}, r_{\bar{D},2}}(t_2), \dots, \widehat{PPV}_{r_{D,J}, r_{\bar{D},J}}(t_J) \right)$, is approximately multivariate normal with,

$$\widehat{PPV}_{r_{D,j}, r_{\bar{D},j}}(t_j) \sim N \left(PPV(t_j), \sigma_{\widehat{PPV}_{r_{D,j}, r_{\bar{D},j}}(t_j)}^2 \right) \quad j=1, 2, \dots, J,$$

$$\sigma^2_{\widehat{PPV}_{r_{D,j},r_{\bar{D},j}}(t_j)} = \left(\frac{t(1-\rho)\rho}{(ROC(t)\rho + t(1-\rho))^2} \right)^2 \sigma^2_{\widehat{ROC}_{r_{D,j},r_{\bar{D},j}}(t_j)},$$

and

$$Cov[\widehat{PPV}_{r_{D,i},r_{\bar{D},i}}(t_i), \widehat{PPV}_{r_{D,j},r_{\bar{D},j}}(t_j)] = \left(\frac{t_i(1-\rho)\rho}{(ROC(t_i)\rho + t_i(1-\rho))^2} \right) \left(\frac{t_j(1-\rho)\rho}{(ROC(t_j)\rho + t_j(1-\rho))^2} \right) Cov[\widehat{ROC}_{r_{D,i},r_{\bar{D},i}}(t_i), \widehat{ROC}_{r_{D,j},r_{\bar{D},j}}(t_j)],$$

where $\sigma^2_{\widehat{ROC}_{r_{D,j},r_{\bar{D},j}}(t_j)}$ and $Cov[\widehat{ROC}_{r_{D,i},r_{\bar{D},i}}(t_i), \widehat{ROC}_{r_{D,j},r_{\bar{D},j}}(t_j)]$ are as defined in Corollary 3.3.

Proof. Immediate from Theorem 3.5.

Corollary A.2

Assume A1-A4 hold and let $\frac{f_D(S^{-1}(t))}{f_{\bar{D}}(S^{-1}(t))}$ be bounded on [a, b]. For $t \in (0,1)$ and J stopping times $(\widehat{PPV}_{r_{D,1},r_{\bar{D},1}}(t), \widehat{PPV}_{r_{D,2},r_{\bar{D},2}}(t), \dots, \widehat{PPV}_{r_{D,J},r_{\bar{D},J}}(t))$, is approximately multivariate normal with,

$$\widehat{PPV}_{r_{D,i},r_{\bar{D},i}}(t) \sim N\left(PPV(t), \sigma^2_{\widehat{PPV}_{r_{D,i},r_{\bar{D},i}}(t)}\right) \quad i=1, 2, \dots, J$$

and

$$Cov[\widehat{PPV}_{r_{D,i},r_{\bar{D},i}}(t), \widehat{PPV}_{r_{D,j},r_{\bar{D},j}}(t)] = var[\widehat{PPV}_{r_{D,j},r_{\bar{D},j}}(t)] = \sigma^2_{\widehat{PPV}_{r_{D,j},r_{\bar{D},j}}(t)}, \quad r_i \leq r_j$$

where $\sigma^2_{\widehat{PPV}_{r_{D,j},r_{\bar{D},j}}(t)}$ is defined as in Corollary A.1.

Proof. Immediate from Corollary A.1.

A.3. Supplementary Results for Section 3.3

Proof of Theorem 3.6

The proof of Theorem 3.6 follows the proofs found in Pyke and Shorack (1968). First, note that,

$$\begin{aligned}
 & n_D^{-1/2}[n_D r_D](\widehat{S}_{D,r_D}(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u)) \\
 & \quad - S_D(F^{-1}(u))) \\
 & = n_D^{-1/2}[n_D r_D](F_D(F^{-1}(u)) - F_D(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u))) \\
 & \quad + n_D^{-1/2}[n_D r_D](F_D(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u)) \\
 & \quad \quad - \widehat{F}_{D,r_D}(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u))).
 \end{aligned}$$

The first term can be rewritten as,

$$\begin{aligned}
 & n_D^{-1/2}[n_D r_D](F_D(F^{-1}(u)) \\
 & \quad - F_D(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u))) > \frac{F_D(F^{-1}(F(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u)))) - F_D(F^{-1}(u))}{F(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u)) - u} \\
 & = n_D^{-1/2}[n_D r_D](F_D(F^{-1}(u)) - F_D(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u))) + \frac{F_D(F^{-1}(F(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u)))) - F_D(F^{-1}(u))}{F(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u)) - u} \rho n_D^{-1/2}[n_D r_D](\widehat{F}_{D,r_D}^{-1}(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u)) \\
 & \quad - F_D(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u))) \\
 & \quad + \frac{n_D^{-1/2}[n_D r_D] F_D(F^{-1}(F(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u)))) - F_D(F^{-1}(u))}{n_D^{-1/2}[n_D r_D] F(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u)) - u} (1 \\
 & \quad - \rho) n_D^{-1/2}[n_D r_{\overline{D}}](\widehat{F}_{D,r_{\overline{D}}}^{-1}(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u)) \\
 & \quad - F_{\overline{D}}(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u)))
 \end{aligned}$$

We begin by showing that $F(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u))$ converges to u uniformly,

$$\begin{aligned}
 & \sup_{c \leq r_D \leq 1} \sup_{d \leq r_{\overline{D}} \leq 1} \sup_{a \leq u \leq b} |F(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u)) \\
 & \quad - u| \leq \sup_{c \leq r_D \leq 1} \sup_{d \leq r_{\overline{D}} \leq 1} \sup_{a \leq u \leq b} |F(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u)) - \widehat{F}_{r_D,r_{\overline{D}}}(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u))| \\
 & \quad + \sup_{c \leq r_D \leq 1} \sup_{d \leq r_{\overline{D}} \leq 1} \sup_{a \leq u \leq b} |F(\widehat{F}_{r_D,r_{\overline{D}}}^{-1}(u)) - u|.
 \end{aligned}$$

We note that,

$$\begin{aligned} & \sup_{c \leq r_D \leq 1} \sup_{d \leq r_{\bar{D}} \leq 1} \sup_{a \leq u \leq b} \left| F\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right) \right. \\ & \quad \left. - \widehat{F}_{r_D, r_{\bar{D}}}\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right) \right| \leq \frac{n_D}{\lfloor n_D c \rfloor} \sup_{c \leq r_D \leq 1} \sup_{d \leq r_{\bar{D}} \leq 1} \sup_{a \leq u \leq b} \frac{\lfloor n_D r_D \rfloor}{n_D} \left| F_D\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right) - \widehat{F}_{r_D, r_{\bar{D}}}\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right) \right| \\ & \quad + \frac{n_{\bar{D}}}{\lfloor n_{\bar{D}} d \rfloor} \sup_{c \leq r_D \leq 1} \sup_{d \leq r_{\bar{D}} \leq 1} \sup_{a \leq u \leq b} \frac{\lfloor n_{\bar{D}} r_{\bar{D}} \rfloor}{n_{\bar{D}}} \left| F_{\bar{D}}\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right) \right. \\ & \quad \left. - F_{\bar{D}, r_{\bar{D}}}\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right) \right| \rightarrow_{a.s.} 0, \end{aligned}$$

by the Glivenko-Cantelli Theorems (1.51 and 1.52 in Csörgő and Szyszkowicz (1998)),

along with the fact that $\frac{n_D}{\lfloor n_D c \rfloor} \rightarrow \frac{1}{c}$ and $\frac{n_{\bar{D}}}{\lfloor n_{\bar{D}} d \rfloor} \rightarrow \frac{1}{d}$. For all $r_D, r_{\bar{D}} \in (0, 1] \times (0, 1]$,

$$\sup_{a \leq u \leq b} \left| u - \widehat{F}_{r_D, r_{\bar{D}}}\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right) \right| \leq_{a.s.} \left(\frac{\rho}{\lfloor r_D n_D \rfloor} \vee \frac{1 - \rho}{\lfloor n_{\bar{D}} r_{\bar{D}} \rfloor} \right).$$

Therefore,

$$\sup_{c \leq r_D \leq 1} \sup_{d \leq r_{\bar{D}} \leq 1} \sup_{a \leq u \leq b} \left| u - \widehat{F}_{r_D, r_{\bar{D}}}\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right) \right| \leq_{a.s.} \left(\frac{\rho}{\lfloor n_D c \rfloor} \vee \frac{1 - \rho}{\lfloor n_{\bar{D}} d \rfloor} \right) \rightarrow 0,$$

which implies that,

$$\sup_{c \leq r_D \leq 1} \sup_{d \leq r_{\bar{D}} \leq 1} \sup_{a \leq u \leq b} \left| F\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right) - u \right| \rightarrow_{a.s.} 0. \quad (\text{A.8})$$

We note that (A.8) also implies that $F_D\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right)$ and $F_{\bar{D}}\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right)$ converge uniformly to $F_D(F^{-1}(u))$ and $F_{\bar{D}}(F^{-1}(u))$, respectively, which can be seen by noting that the difference between $F_D\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right)$ and $F_D(F^{-1}(u))$ will always have the same sign as the difference between $F_{\bar{D}}\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right)$ and $F_{\bar{D}}(F^{-1}(u))$.

By the mean value theorem, there exists $F\left(\widetilde{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right)$ between u and $F\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right)$ such that,

$$\frac{F_D\left(F^{-1}\left(F\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right)\right)\right) - F_D\left(F^{-1}(u)\right)}{F\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right) - u} = \frac{f_D\left(F^{-1}\left(F\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right)\right)\right)}{f\left(F^{-1}\left(F\left(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)\right)\right)\right)}.$$

The uniform continuity of $\frac{f_D(F^{-1}(u))}{f(F^{-1}(u))}$ combined with the fact that $F(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)) \rightarrow_{a.s.} u$ uniformly, allows us to conclude,

$$\sup_{c \leq r_D \leq 1} \sup_{d \leq r_{\bar{D}} \leq 1} \sup_{a \leq u \leq b} \left| \frac{f_D(F^{-1}(F(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u))))}{f(F^{-1}(F(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u))))} - \frac{f_D(F^{-1}(u))}{f(F^{-1}(u))} \right| \rightarrow_{a.s.} 0. \quad (A.9)$$

For all $r_D, r_{\bar{D}} \in (0, 1] \times (0, 1]$,

$$\sup_{a \leq u \leq b} n_D^{-1/2} [n_D r_D] \left| u - \widehat{F}_{r_D, r_{\bar{D}}}(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)) \right| \leq_{a.s.} \left(\frac{\rho}{n_D^{-1/2}} \vee \frac{[n_D r_D] (1 - \rho)}{[n_{\bar{D}} r_{\bar{D}}] n_D^{-1/2}} \right).$$

Therefore, as $n_D \rightarrow \infty$ and $n_{\bar{D}} \rightarrow \infty$,

$$\sup_{0 \leq r_D \leq 1} \sup_{0 \leq r_{\bar{D}} \leq 1} \sup_{a \leq u \leq b} n_D^{-1/2} [n_D r_D] \left| u - \widehat{F}_{r_D, r_{\bar{D}}}(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)) \right| \rightarrow_{a.s.} 0.$$

Combining this result with (A.9) allows us to conclude that,

$$\frac{F_D(F^{-1}(F(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)))) - F_D(F^{-1}(u))}{F(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)) - u} \rightarrow_{a.s.} n_D^{-1/2} [n_D r_D] (u - \widehat{F}_{r_D, r_{\bar{D}}}(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u))) \rightarrow_{a.s.} 0.$$

Corollary 1.A in Csörgő and Szyszkowicz (1998), (A.9) and the uniform continuity of the Kiefer process allow us to conclude,

$$\begin{aligned} & \frac{F_D(F^{-1}(F(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)))) - F_D(F^{-1}(u))}{F(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)) - u} \rightarrow_d \rho n_D^{-1/2} [n_D r_D] \left(\widehat{F}_{D, r_D}(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)) \right) \\ & - F_D(\widehat{F}_{r_D, r_{\bar{D}}}^{-1}(u)) \rightarrow_d \frac{f_D(F^{-1}(u))}{f(F^{-1}(u))} \rho K_1(F_D(F^{-1}(u)), r_D), \end{aligned} \quad (A.10)$$

and,

$$\begin{aligned}
 & \frac{n_D^{-1/2}[n_D r_D] F_D \left(F^{-1} \left(F \left(\widehat{F}_{r_D, r_{\overline{D}}}^{-1}(u) \right) \right) \right) - F_D \left(F^{-1}(u) \right)}{n_D^{-1/2} \left[n_{\overline{D}} r_{\overline{D}} \right]} (1) \\
 & - \rho) n_D^{-1/2} \left[n_{\overline{D}} r_{\overline{D}} \right] \left(\widehat{F}_{r_D, r_{\overline{D}}}^{-1} \left(\widehat{F}_{r_D, r_{\overline{D}}}^{-1}(u) \right) \right) \quad (A.11) \\
 & - F_{\overline{D}} \left(\widehat{F}_{r_D, r_{\overline{D}}}^{-1}(u) \right) \rightarrow_d \sqrt{\lambda} \frac{r_D}{r_{\overline{D}}} \frac{f_D \left(F^{-1}(u) \right)}{f \left(F^{-1}(u) \right)} (1) \\
 & - \rho) K_2 \left(F_{\overline{D}} \left(F^{-1}(u) \right), r_D \right).
 \end{aligned}$$

The second term converges in distribution to a Kiefer process,

$$\begin{aligned}
 & n_D^{-1/2} [n_D r_D] \left(F_D \left(\widehat{F}_{r_D, r_{\overline{D}}}^{-1}(u) \right) - \widehat{F}_{D, r_D} \left(\widehat{F}_{r_D, r_{\overline{D}}}^{-1}(u) \right) \right) = -n_D^{-1/2} [n_D r_D] \left(\widehat{F}_{D, r_D} \left(\widehat{F}_{r_D, r_{\overline{D}}}^{-1}(u) \right) \right. \\
 & \left. - F_D \left(\widehat{F}_{r_D, r_{\overline{D}}}^{-1}(u) \right) \right) \rightarrow_d \quad (A.12) \\
 & - K_1 \left(F_D \left(F^{-1}(u) \right), r_D \right),
 \end{aligned}$$

by Corollary 1.A in Csörgő and Szyszkowicz (1998). Summing (A.10), (A.11) and (A.12) gives the desired result.

Corollary A.3

Assume A1 - A4 hold and let $\frac{f_D \left(F^{-1}(u) \right)}{f \left(F^{-1}(u) \right)}$ be bounded on $[a, b]$. For $u_1, u_2, \dots, u_J \in (0, 1)$, $r_{D,1}, r_{D,2}, \dots, r_{D,J} \in (0, 1]$ and $r_{\overline{D},1}, r_{\overline{D},2}, \dots, r_{\overline{D},J} \in (0, 1]$, a vector of arbitrary points on the sequential empirical PPV curve, $\left(\widehat{PPV}_{r_{D,1}, r_{\overline{D},1}}(u_1), \widehat{PPV}_{r_{D,2}, r_{\overline{D},2}}(u_2), \dots, \widehat{PPV}_{r_{D,J}, r_{\overline{D},J}}(u_J) \right)$, is approximately multivariate normal with,

$$\widehat{PPV}_{r_{D,j}, r_{\overline{D},j}}(u_j) \sim N \left(PPV(u_j), \sigma_{\widehat{PPV}_{r_{D,j}, r_{\overline{D},j}}(u_j)}^2 \right) \quad j=1, 2, \dots, J$$

with,

$$\begin{aligned}
 & Cov \left[\widehat{PPV}_{r_{D,1}, r_{\overline{D},1}}(u_1), \widehat{PPV}_{r_{D,2}, r_{\overline{D},2}}(u_2) \right] \\
 & = \frac{(1-\rho)^2 u_1}{(1-u_1)} \frac{f_{\overline{D}} \left(F^{-1}(u_1) \right)}{f \left(F^{-1}(u_1) \right)} \frac{f_{\overline{D}} \left(F^{-1}(u_2) \right)}{f \left(F^{-1}(u_2) \right)} \frac{(r_{D,1} \wedge r_{D,2}) (1 - NPV(u_1)) PPV(u_2)}{n_D r_{D,1} r_{D,2}} \\
 & + \frac{\rho^2 u_1}{(1-u_1)} \frac{f_D \left(F^{-1}(u_1) \right)}{f \left(F^{-1}(u_1) \right)} \frac{f_D \left(F^{-1}(u_2) \right)}{f \left(F^{-1}(u_2) \right)} \frac{(r_{\overline{D},2} \wedge r_{\overline{D},1}) NPV(u_1) (1 - PPV(u_2))}{n_D r_{D,1} r_{D,2}},
 \end{aligned}$$

when $u_1 \neq u_2$ and,

$$\begin{aligned}
 & Cov \left[\widehat{PPV}_{r_{D,1}, r_{\bar{D},1}}(u_1), \widehat{PPV}_{r_{D,2}, r_{\bar{D},2}}(u_2) \right] \\
 &= \frac{(1-\rho)^2 u_2}{(1-u_2)} \frac{f_D(F^{-1}(u_1))}{f(F^{-1}(u_1))} \frac{f_{\bar{D}}(F^{-1}(u_2))}{f(F^{-1}(u_2))} \frac{(r_{D,1} \wedge r_{D,2})(1-NPV(u_2))PPV(u_1)}{n_D r_{D,1} r_{D,2}} \\
 &+ \frac{\rho^2 u_2}{(1-u_2)} \frac{f_D(F^{-1}(u_1))}{f(F^{-1}(u_1))} \frac{f_{\bar{D}}(F^{-1}(u_2))}{f(F^{-1}(u_2))} \frac{(r_{\bar{D},2} \wedge r_{\bar{D},2})NPV(u_2)(1-PPV(u_1))}{n_D r_{D,1} r_{D,2}},
 \end{aligned}$$

when $u_2 \geq u_1$, where $\sigma^2_{\widehat{PPV}_{r_{D,1}, r_{\bar{D},1}}(u_1)}$ is defined as in Corollary 3.8.

Proof. Immediate from Theorem 3.6.

References

Csörgő, M.; Szyszkowicz, B. Order statistics: theory & methods Handbook of Statist. Vol. 16. North-Holland; Amsterdam: 1998. Sequential quantile and Bahadur-Kiefer processes; p. 631-688. MR1668760

Dorfman D, Alf E. Maximum likelihood estimation of parameters of signal detection theory and determination of confidence intervals - rating method data. *Journal of Mathematical Psychology.* 1960; 6:487-496.

Hsieh F, Turnbull WB. Nonparametric and semiparametric estimation of the receiver operating characteristic curve. *The Annals of Statistics.* 1996; 24:25-40.

Hwang IK, Shih WJ, De Cani JS. Group sequential designs using a family of type I error probability spending functions. *Statistics in Medicine.* 1990; 9:1439-1445. [PubMed: 2281231]

Jennison, C.; Turnbull, BW. Group sequential methods with applications to clinical trials. CRC Press Inc; 2000.

Koopmeiners JS, Feng Z. Asymptotic Properties of the Sequential Empirical ROC and PPV Curves Working Paper report No 345. 2010

Liu A, Hall WJ. Unbiased estimation of secondary parameters following a sequential test. *Biometrika.* 2001; 88:895-900.

Marrero JA, Feng Z, Wang Y, Nguyen MH, Befeler AS, Roberts LR, Reddy KR, Harnois D, Llovet JM, Normolle D, Dalhgren J, Chia D, Lok AS, Wagner PD, Srivastava S, Schwartz M. [alpha]-Fetoprotein, Des-[gamma] Carboxyprothrombin, and Lectin-Bound [alpha]-Fetoprotein in Early Hepatocellular Carcinoma. *Gastroenterology.* 2009; 137:110-118. [PubMed: 19362088]

Moskowitz CS, Pepe MS. Quantifying and comparing the predictive accuracy of continuous prognostic factors for binary outcomes. *Biostatistics.* 2004; 5:113-127. [PubMed: 14744831]

Pepe, MS. The statistical evaluation of medical tests for classification and prediction Oxford Statistical Science Series. Vol. 28. Oxford University Press; Oxford: 2003. MR2260483 (2008f:62006)

Pepe MS, Feng Z, Longton G, Koopmeiners J. Conditional Estimation of Sensitivity and Specificity from a Phase 2 Biomarker Study Allowing Early Termination for Futility. *Statistics in Medicine.* 2009; 28:762-779. [PubMed: 19097251]

Pyke R, Shorack G. Weak Convergence of a Two-sample Empirical Process and a New Approach to Chernoff-Savage Theorems (Corr: 75AnlsProb 3 P1068). *The Annals of Mathematical Statistics.* 1968; 39:755-771.

Tang L, Emerson SS, Zhou XH. Nonparametric and Semiparametric Group Sequential Methods for Comparing Accuracy of Diagnostic Tests. *Biometrics.* 2008; 64:1137-1145. [PubMed: 18371124]

Tang LL, Liu A. Sample size recalculation in sequential diagnostic trials. *Biostatistics.* 2010; 11:151-163. [PubMed: 19822693]

Zheng Y, Cai T, Pepe MS, Levy WC. Time-Dependent Predictive Values of Prognostic Biomarkers With Failure Time Outcome. *Journal of the American Statistical Association.* 2008; 103:362-368. [PubMed: 19655041]

Zhu H, Hu F. Sequential monitoring of response-adaptive randomized clinical trials. *The Annals of Statistics*. 2010; 38:2218–2241.

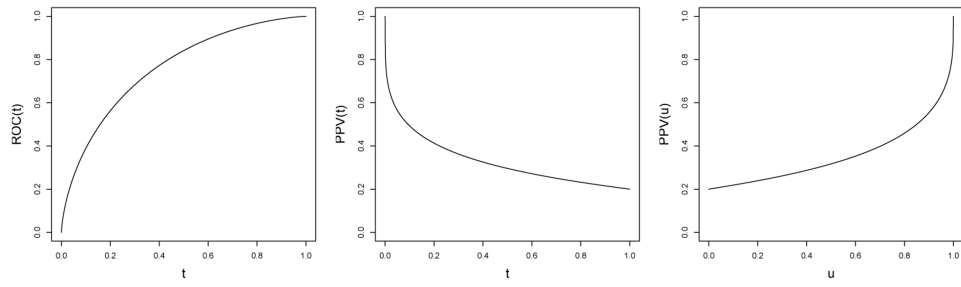


Fig 1.
True ROC and PPV curves for the scenario considered in Section 4.

Table 1

Simulation results to evaluate the finite sample properties of Theorem 3.1. Presented are the expected value, simulated probability of being less than 5th, 25th, 50th, 75th and 95th percentile of the normal distribution, the simulated covariance matrix and the theoretical covariance matrix for R_{r_D, r_D} (t). 10000 simulations were performed for each scenario.

	Observed						Theoretical							
	Mean	5th %tile	25th %tile	50th %tile	75th %tile	95th %tile	Covariance Matrix							
	$n_D = 50$													
$R_{0.4,0.7}(0.4)$	0.01	0.05	0.17	0.46	0.63	0.98	0.1	0.117	0.079	0.103	0.104	0.129	0.081	0.104
$R_{1,1}(0.4)$	0.02	0.07	0.2	0.44	0.74	0.97		0.318	0.104	0.262		0.322	0.104	0.26
$R_{0.4,0.7}(0.2)$	0.03	0.04	0.22	0.47	0.73	0.96			0.161	0.201			0.171	0.225
$R_{1,1}(0.2)$	0.05	0.04	0.2	0.47	0.68	0.93				0.544				0.563
	$n_D = 100$													
$R_{0.4,0.7}(0.4)$	0.01	0.05	0.21	0.41	0.78	0.97	0.101	0.12	0.08	0.102	0.104	0.129	0.081	0.104
$R_{1,1}(0.4)$	0.02	0.05	0.24	0.48	0.76	0.96		0.318	0.104	0.26		0.322	0.104	0.26
$R_{0.4,0.7}(0.2)$	0.03	0.04	0.2	0.45	0.73	0.95			0.164	0.205			0.171	0.225
$R_{1,1}(0.2)$	0.05	0.04	0.23	0.47	0.73	0.95				0.55				0.563
	$n_D = 200$													
$R_{0.4,0.7}(0.4)$	0.01	0.06	0.22	0.44	0.7	0.96	0.104	0.121	0.081	0.102	0.104	0.129	0.081	0.104
$R_{1,1}(0.4)$	0.02	0.05	0.25	0.48	0.72	0.95		0.317	0.104	0.259		0.322	0.104	0.26
$R_{0.4,0.7}(0.2)$	0.03	0.04	0.25	0.5	0.7	0.94			0.168	0.212			0.171	0.225
$R_{1,1}(0.2)$	0.05	0.05	0.23	0.46	0.72	0.95				0.555				0.563

Table 2

Simulation results to evaluate the operating characteristics of a study to evaluate the predictive accuracy of DCP using a fixed-sample design and group sequential designs with two, three or four stopping times. Presented are the probability of rejecting the null hypothesis and expected sample size under the null and alternative hypotheses. 10000 simulations were performed for each scenario.

Stopping Times	$NPV(0.6) = 0.90$		$NPV(0.6) = 0.95$		$NPV(0.6) = 0.90$		$NPV(0.6) = 0.95$	
	$P(reject)$	$E(n_D)$	$P(reject)$	$E(n_D)$	$P(reject)$	$E(n_D)$	$P(reject)$	$E(n_D)$
J = 1	0.003	702	0.03	702	0.026	702	0.917	702
J = 2	0.004	432	0.026	492.4	0.024	489.5	0.924	624.5
J = 3	0.004	367.4	0.022	431.3	0.023	433	0.917	580.1
J = 4	0.002	340	0.023	410.7	0.024	417.2	0.911	571.1