2-28-2013

Asymptotic and Finite Sample Behavior of Net Reclassification Indices

Zheyu Wang
University of Washington - Seattle Campus, wangzy@u.washington.edu

Asymptotic and Finite Sample Behavior of Net Reclassification Indices

Zheyu Wang

Abstract

The Net Reclassification Index (NRI) introduced by Pencina and colleagues [1, 2] is designed to quantify the prediction increment provided by a new biomarker. It has become popular for evaluating and selecting novel markers. The published variance formulae for NRI statistics do not account for the fact that risks are estimated based on risk models fit to data, and thus are not valid in practice when estimated risks are used [3]. Kerr and colleagues [4] showed that the confidence intervals constructed based on a bootstrap estimate of the variance and Normal approximation had the best performance among various methods they examined, including the one based on bootstrap quantiles. This paper establishes asymptotic Normality of NRI statistics when true risks are unknown and are estimated. Our results provide theoretical support for constructing confidence intervals for NRI statistics based on a Normal approximation. We also derive explicit variance formulae for NRI statistics that are calculated based on estimated risks. In addition, we examine finite sample distributional behavior of NRI statistics in a simulation study. These results provide some guidance on the sample size required for adopting a Normal approximation for NRI inference in practice.

Keyworks: Net reclassification index; incremental value; asymptotic Normality; sampling distribution; biomarker; risk models

1 Introduction

Risk prediction is key in medical decision making as well as in health policy development. Accurate risk prediction can assist clinicians in recommending the most beneficial treatment to patients while avoiding unnecessary, invasive, or costly procedures. It is also crucial for health policy makers to develop well-informed strategies for the whole population. This has led to continuous efforts to improve prediction models and successive discovery of novel markers. For example, breast density [5, 6] and genetic polymorphisms [7, 8, 9] are proposed for predicting breast cancer risk in addition to traditional factors in the Gail model [10, 11]. Numerous studies have been conducted in recent years to evaluate candidate markers for cardiovascular event upon factors in the standard Framingham risk score [12]. An important question is how best to assess and quantify the improvement gained from incorporating new biomarkers into risk prediction models.

Various metrics have been proposed to quantify the prediction increment, or incremental value, of a biomarker [13]. Change in the area under the receiver operating characteristic curve (ΔAUC) is the most popular single number summary index. However, AUC has been criticized because it does not measure a clinically meaningful quantity, and because it is a broad summary of changes in risk models that incorporates irrelevant information [14, 15, 16]. To overcome these limitations, Pencina and colleagues [1] proposed the Net Reclassification Index (NRI) as a new measure of incremental value. The original definition of NRI is based on a reclassification table with predefined risk categories. It is conceived from the idea that a useful biomarker will lead to more diseased subjects in higher risk categories and more healthy subjects in lower risk categories. Consequently, it contains two parts, the "event"

NRI, and the "nonevent" NRI, as follows,

$$\begin{aligned} NRI(event) &= P(up|event) - P(down|event), \\ NRI(nonevent) &= P(down|nonevent) - P(up|nonevent), \\ NRI &= NRI(event) + NRI(nonevent). \end{aligned}$$

Here, "event" denotes subjects with disease or other medical conditions of interest ("cases") and "nonevent" denotes controls. "Up" indicates that the risk predication based on the model with the new biomarker moves an individual into a higher risk category compared to the old model with baseline predictors. "Down" indicates the reverse, that the risk predication based on the new model moves an individual into a lower risk category compared to the old model. Later, a "category-free" or "continuous" NRI was introduced [2] to avoid the need for subjective and perhaps arbitrary risk thresholds, although this version of NRI shares many of the same limitations as ΔAUC [4, 17]. Despite their limitations, NRI statistics have become increasing popular, especially in cardiovascular research [18, 19].

A close examination of the asymptotic behavior of NRI statistics is necessary to correctly gauge the uncertainty in the estimation and construct valid confidence intervals for inference. Pencina and colleagues provided formulae [2] for NRI statistics comparing risks calculated from the baseline risk model and risks calculated from the expanded risk model that includes the new biomarker. Pepe et al. [3] pointed out that the variance formulae derived based on fixed risk models, such as the ones in Pencina et al. [2], do not consider the variability in regression model coefficient estimates, and thus are not valid when estimated risks are used. Kerr et al. [4] illustrated this issue for NRI statistics with a simulation study. In practice, the risk model is rarely known and is almost always estimated from the data. Therefore, it is of interest to derive variance formulae for NRI statistics that account for the variability in risk estimates and can be used in practical situations. Similarly, it is critical to examine the asymptotic and finite sample distributional behavior of NRI statistics, since the commonly

used method to construct a 95% confidence interval for a parameter θ, $\hat{\theta} \pm 1.96 \cdot \widehat{SE}(\hat{\theta})$, requires the distribution of $\hat{\theta}$ to be approximately Normal, regardless of whether the standard error estimate $\widehat{SE}(\hat{\theta})$ is obtained from a formula or by bootstrapping. The latter choice, confidence interval for NRI statistics constructed based on a bootstrapped standard error and Normal approximation, exhibited better coverage performance in the simulation study in Kerr et al. [4] than various other confidence intervals they examined.

In this paper, we derive the asymptotic distribution of NRI statistics when the estimated risks are used. We also study their finite sample behavior with simulations. Based on the recommendation in Kerr et al. [4], we focuse on the category-free NRI and the two-category NRI. This paper is organized as follows. Section 2 describes notation, settings and assumptions. Section 3 and section 4 provide the derivations of the asymptotic distributions of the category-free NRI and the two-category NRI, respectively. Section 5 studies the finite sample behavior of NRI statistics via simulations. Section 6 concludes the paper with a discussion.

2 Notation and settings

Suppose that we have a set of baseline risk factors $\mathbf{X}$ and a new biomarker Y. The task is to evaluate the prediction increment introduced by this new biomaker. To do this, we want to compare the classification performance of a new model with both $\mathbf{X}$ and Y as predictors with that of the old model with only baseline predictors $\mathbf{X}$.

Let $F_0 = F_{\mathbf{X}}$ be the distribution function of $\mathbf{X}$, and $F_1 = F_{\mathbf{X},Y}$ be the joint distribution of $\mathbf{X}$ and Y. Let D be a binary variable indicate subject's disease status, i.e., D equals 1 for cases and D equals 0 for controls. Suppose $R_0(\beta, \mathbf{X}) = P(D = 1|\mathbf{X})$, $R_1(\theta, \mathbf{X}, Y) = P(D = 1|\mathbf{X}, Y)$ are the old and new risk models, respectively. We further suppose that the

assumed risk models are the true risk models, i.e., there is no model misspecification. And $\hat{F}_0$, $\hat{F}_1$, $\hat{R}_0 = R_0(\hat{\beta}, \mathbf{X})$ and $\hat{R}_1 = R_1(\hat{\theta}, \mathbf{X}, Y)$ are the corresponding estimated quantities. In addition, n denotes the total sample size, and n_D denotes the number of cases.

Assumptions:

(1) $\sqrt{n}(\hat{\beta} - \beta) \to_d N(0, \Sigma_0(\beta))$ and $\sqrt{n}(\hat{\theta} - \theta) \to_d N(0, \Sigma_1(\theta))$.

(2) Risk functions $R_0(r, \mathbf{X})$ and $R_1(s, \mathbf{X}, Y)$ are differentiable at true value $r = \beta$ and $s = \theta$.

(3) $R_0^{-1}(r, t)$ and $R_1^{-1}(s, t)$ exist and are differentiable at $r = \beta$ and $s = \theta$.

(4) The set $\Omega = \{(\mathbf{X}, Y) | R_0(r, \mathbf{X}) \neq R_1(s, \mathbf{X}, Y)\}$ has positive measure.

(5) F_0 and F_1 are continuous with positive density f_0 and f_1 on [0, 1] except on finite points.

(6) $n_D/n \to_p \rho > 0$, as $n \to +\infty$.

3 Asymptotic Distribution of a Category-Free NRI

A category-free NRI, also called continuous NRI, is the summation of an event NRI and a nonevent NRI that are calculated based on continuous risks without pre-selected threshold. Event NRI is defined as the probability that the new risk model provides a higher risk than the old model among cases, $P\big(R_1(\mathbf{X}, Y) > R_0(\mathbf{X}) | D = 1\big)$, minus the probability that the new risk model provides a lower risk than the old model among cases, $P\big(R_1(\mathbf{X}, Y) < R_0(\mathbf{X}) | D = 1\big)$. Nonevent NRI is defined as the probability that the new risk model provides a lower risk than the old model among controls, $P\big(R_1(\mathbf{X}, Y) < R_0(\mathbf{X}) | D = 0\big)$, minus the probability that the new risk model provides a higher risk than the old model among controls, $P\big(R_1(\mathbf{X}, Y) > R_0(\mathbf{X}) | D = 0\big)$.

In most situations, risks need to be estimated in addition to the four probabilities comparing

them. So the estimated category-free NRI is calculated as follows,

$$\begin{aligned}
\widehat{NRI_e^{>0}} &= \hat{P}\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big) - \hat{P}\big(\hat{R}_1(\mathbf{X},Y) < \hat{R}_0(\mathbf{X})|D=1\big), \\
\widehat{NRI_{ne}^{>0}} &= \hat{P}\big(\hat{R}_1(\mathbf{X},Y) < \hat{R}_0(\mathbf{X})|D=0\big) - \hat{P}\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=0\big), \\
\widehat{NRI^{>0}} &= \widehat{NRI_e^{>0}} + \widehat{NRI_{ne}^{>0}}.
\end{aligned}$$

$NRI_e^{>0}$ characterizes the reclassification improvement among disease population, and $NRI_{ne}^{>0}$ summarizes the improvement among non-diseased population. Most often the costs and benefits of these improvements will differ greatly for cases and controls. Therefore, it is recommended to report these two components of the NRI statistic separately [4]. We will derive the asymptotic results separately for $\widehat{NRI_e^{>0}}$ and $\widehat{NRI_{ne}^{>0}}$.

3.1

We first consider $\widehat{NRI_e^{>0}}$.

$$\begin{aligned}
\widehat{NRI_e^{>0}} &= \hat{P}\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big) - P\big(\hat{R}_1(\mathbf{X},Y) < \hat{R}_0(\mathbf{X})|D=1\big) \\
&= \hat{P}\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big) - \big(1 - \hat{P}\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big)\big) \\
&\quad \text{(By assumption 2)} \\
&= 2\hat{P}\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big) - 1.
\end{aligned}$$

$$\begin{aligned}
&\sqrt{n}\big[\hat{P}\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big) - P\big(R_1(\mathbf{X},Y) > R_0(\mathbf{X})|D=1\big)\big] \\
= &\Big\{\sqrt{n}\big[\hat{P}\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big) - P\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big)\big]\Big\} \\
&+\Big\{\sqrt{n}\big[P\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big) - P\big(R_1(\mathbf{X},Y) > R_0(\mathbf{X})|D=1\big)\big]\Big\} \\
= &\big\{A\big\} + \big\{B\big\}.
\end{aligned}$$

Given $\{B\}$, or equivalently $P\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big)$, the first term $\{A\}$ is the empirical value of B derived from a proportion among the n_D events. Therefore, conditioning on B, A has mean 0 and binomial variance

$$\begin{aligned} var(A|B) &= nP\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big)\Big[1 - P\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big)\Big]/n_D \\ &\doteq \frac{1+NRI_e^{>0}}{2\rho} \times \frac{1-NRI_e^{>0}}{2} = \frac{1-(NRI_e^{>0})^2}{4\rho}. \end{aligned}$$

This can also be seen by noting that,

$$\begin{aligned} & \sqrt{n}\big[\hat{P}\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big) - \hat{P}\big(R_1(\mathbf{X},Y) > R_0(\mathbf{X})|D=1\big)\big] \\ - \; & \sqrt{n}\big[P\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big) - P\big(R_1(\mathbf{X},Y) > R_0(\mathbf{X})|D=1\big)\big] = o_p(1). \\ & \text{(By equicontinuity of process } \sqrt{n_D}(\hat{P}-P) \text{ and assumption 1, 2.)} \end{aligned}$$

Thus, $$A = \sqrt{n}\big[\hat{P}\big(R_1(\mathbf{X},Y) > R_0(\mathbf{X})|D=1\big) - P\big(R_1(\mathbf{X},Y) > R_0(\mathbf{X})|D=1\big)\big] + o_p(1).$$

Because $E(A|B)=0$, we have that $var(A) = E\big[var(A|B)\big] \doteq [1-(NRI_e^{>0})^2]/4\rho$. Moreover, $E(A|B)=0$ also implies A and B are uncorrelated. Hence,

$$\begin{aligned} & var\big(\sqrt{n}\widehat{NRI_e^{>0}}\big) \\ = \; & 4var\Big\{\sqrt{n}\big[\hat{P}\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big) - P\big(R_1(\mathbf{X},Y) > R_0(\mathbf{X})|D=1\big)\big]\Big\} \\ = \; & 4\big\{var(A) + var(B)\big\}. \end{aligned}$$

3.2

Now we turn to the variance of $B = \sqrt{n}\big[P\big(\hat{R}_1(\mathbf{X},Y) > \hat{R}_0(\mathbf{X})|D=1\big) - P\big(R_1(\mathbf{X},Y) > R_0(\mathbf{X})|D=1\big)\big]$. The idea is to write it as a function of β and θ based on the asymptotic distribution of $\hat{R}_1(\mathbf{X},Y) - \hat{R}_0(\mathbf{X})$, and then apply the delta method.

First, we derive asymptotic distribution of $\hat{R}_1(\mathbf{X}, Y) - \hat{R}_0(\mathbf{X})$ by Taylor expansion and assumption 1.

$$\begin{aligned} & \sqrt{n}\big\{\big[\hat{R}_1(\mathbf{X}, Y) - \hat{R}_0(\mathbf{X})\big] - \big[R_1(\mathbf{X}, Y) - R_0(\mathbf{X})\big]\big\} \\ = \; & \big(\frac{\partial R_1(s, \mathbf{X}, Y)}{\partial s}|_{s=\theta}\big)^T \sqrt{n}(\hat{\theta} - \theta) - \big(\frac{\partial R_0(r, \mathbf{X})}{\partial r}|_{r=\beta}\big)^T \sqrt{n}(\hat{\beta} - \beta) + o_p(1). \end{aligned}$$

Thus, $\sqrt{n}\big\{\big[\hat{R}_1(\mathbf{X}, Y) - \hat{R}_0(\mathbf{X})\big] - \big[R_1(\mathbf{X}, Y) - R_0(\mathbf{X})\big]\big\} \rightarrow_d N(0, \Sigma_R)$ by Slutskys theorem, where

$$\begin{aligned} \Sigma_R \; = \; & \big(\frac{\partial R_0(r, \mathbf{X})}{\partial r}|_{r=\beta}\big)^T \Sigma_0(\beta) \big(\frac{\partial R_0(r, \mathbf{X})}{\partial r}|_{r=\beta}\big) + \big(\frac{\partial R_1(s, \mathbf{X}, Y)}{\partial s}|_{s=\theta}\big)^T \Sigma_1(\theta) \big(\frac{\partial R_1(s, \mathbf{X}, Y)}{\partial s}|_{s=\theta}\big) \\ & -2\big(\frac{\partial R_0(r, \mathbf{X})}{\partial r}|_{r=\beta}\big)^T cov\big(\sqrt{n}(\hat{\beta} - \beta), \sqrt{n}(\hat{\theta} - \theta)\big) \big(\frac{\partial R_1(s, \mathbf{X}, Y)}{\partial s}|_{s=\theta}\big). \end{aligned}$$

In the above, the random variable $\sqrt{n}\big\{\big[\hat{R}_1(\mathbf{X}, Y) - \hat{R}_0(\mathbf{X})\big] - \big[R_1(\mathbf{X}, Y) - R_0(\mathbf{X})\big]\big\}$ is not degenerate due to assumption 4. Nevertheless, when assumption 4 does not hold, that is, $R_0(r, \mathbf{X}) = R_1(s, \mathbf{X}, Y), a.s.$, we have,

$$var(A) \doteq \frac{1 - (NRI_e^{>0})^2}{4\rho} = \frac{1}{4\rho} \neq 0.$$

Thus, $var\big(\sqrt{n}\widehat{NRI_e^{>0}}\big) = 4\big\{var(A) + var(B)\big\}$ is dominated by $var(A)$. In fact,

$$var\big(\sqrt{n}\widehat{NRI_e^{>0}}\big) = 4var(A) = \frac{1}{\rho} \neq 0.$$

So the derivation can still go through. Thus, assumption 4 is not essential in establishing asymptotic Normality of $\widehat{NRI_e^{>0}}$, and similarly of $\widehat{NRI_{ne}^{>0}}$. However, it is required in the derivations for the two-category NRI as we will see in section 4.

Let ψ_{0i} and ψ_{1i} be the influence functions of $\hat{\beta}$ and $\hat{\theta}$, respectively, such that,

$$\sqrt{n}(\hat{\beta}-\beta)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_{0i}(\mathbf{X})+o_p(1),\quad \sqrt{n}(\hat{\theta}-\theta)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_{1i}(\mathbf{X},Y)+o_p(1).$$

Then,

$$\begin{aligned}\Sigma_R &= \big(\frac{\partial R_0(\beta,\mathbf{X})}{\partial\beta}\big)^T var(\psi_{0i}(\mathbf{X}))\big(\frac{\partial R_0(\beta,\mathbf{X})}{\partial\beta}\big)+\big(\frac{\partial R_1(\theta,\mathbf{X},Y)}{\partial\theta}\big)^T var(\psi_{1i}(\mathbf{X},Y))\big(\frac{\partial R_1(\theta,\mathbf{X},Y)}{\partial\theta}\big)\\ &\quad -2\big(\frac{\partial R_0(\beta,\mathbf{X})}{\partial\beta}\big)^T cov\big(\psi_{0i}(\mathbf{X}),\psi_{1i}(\mathbf{X},Y)\big)\big(\frac{\partial R_1(\theta,\mathbf{X},Y)}{\partial\theta}\big).\end{aligned}$$

Therefore, we have the following result asymptotically:

$$\begin{aligned}&P\big(\hat{R}_1(\mathbf{X},Y)>\hat{R}_0(\mathbf{X})|D=1\big)=P\big(\hat{R}_1(\mathbf{X},Y)-\hat{R}_0(\mathbf{X})>0|D=1\big)\\ \doteq\ &\int 1-\Phi\big(\frac{0-\big[R_1(\theta,\mathbf{X},Y)-R_0(\beta,\mathbf{X})\big]}{\Sigma_R^{1/2}(\beta,\theta)}\big)dF_{1D}(\mathbf{X},Y)\\ =\ &\int\Phi\big(\frac{\big[R_1(\theta,\mathbf{X},Y)-R_0(\beta,\mathbf{X})\big]}{\Sigma_R^{1/2}(\beta,\theta)}\big)dF_{1D}(\mathbf{X},Y)\equiv H(\beta,\theta),\\ &\text{where }\Phi(.)\text{ is the cdf of a standard Normal variable.}\end{aligned}$$

By Taylor expansion we have,

$$\begin{aligned}B &= \sqrt{n}\big[P\big(\hat{R}_1(\mathbf{X},Y)>\hat{R}_0(\mathbf{X})|D=1\big)-P\big(R_1(\mathbf{X},Y)>R_0(\mathbf{X})|D=1\big)\big]\\ &= \big(\frac{\partial H(r,s)}{\partial r}|_{r=\beta,s=\theta}\big)^T\sqrt{n}(\hat{\beta}-\beta)+\big(\frac{\partial H(r,s)}{\partial s}|_{s=\beta,t=\theta}\big)^T\sqrt{n}(\hat{\theta}-\theta)+o_p(1).\end{aligned}$$

Thus, B is asymptotically Normal with mean 0 and variance

$$\begin{aligned}var(B) &= \big(\frac{\partial H(\beta,\theta)}{\partial\beta}\big)^T\Sigma_0(\beta)\big(\frac{\partial H(\beta,\theta)}{\partial\beta}\big)+\big(\frac{\partial H(\beta,\theta)}{\partial\theta}\big)^T\Sigma_1(\theta)\big(\frac{\partial H(\beta,\theta)}{\partial\theta}\big)\\ &\quad +2\big(\frac{\partial H(\beta,\theta)}{\partial\beta}\big)^T cov\big(\sqrt{n}(\hat{\beta}-\beta),\sqrt{n}(\hat{\theta}-\theta)\big)\big(\frac{\partial H(\beta,\theta)}{\partial\theta}\big).\end{aligned}$$

3.3

We have shown that A and B are two uncorrelated and asymptotically Normal random variables. Thus,

$$\sqrt{n}\big[\widehat{NRI_e^{>0}} - NRI_e^{>0}\big] \to_d N(0, \Sigma_{NRI_e^{>0}}).$$

The asymptotic zero mean of $\sqrt{n}\big[\widehat{NRI_e^{>0}} - NRI_e^{>0}\big]$ is because of equicontinuity of process $\sqrt{n_D}(\hat{P} - P)$ and assumptions 1 and 2. Moreover,

$$\begin{aligned}
&\Sigma_{NRI_e^{>0}} = var\big(\sqrt{n}\widehat{NRI_e^{>0}}\big) = 4\big\{var(A) + var(B)\big\} \\
= \;& \frac{1-(NRI_e^{>0})^2}{\rho} + 4\big(\frac{\partial H(\beta,\theta)}{\partial \beta}\big)^T \Sigma_0(\beta)\big(\frac{\partial H(\beta,\theta)}{\partial \beta}\big) + 4\big(\frac{\partial H(\beta,\theta)}{\partial \theta}\big)^T \Sigma_1(\theta)\big(\frac{\partial H(\beta,\theta)}{\partial \theta}\big) \\
& +8\big(\frac{\partial H(\beta,\theta)}{\partial \beta}\big)^T cov\big(\sqrt{n}(\hat{\beta}-\beta), \sqrt{n}(\hat{\theta}-\theta)\big)\big(\frac{\partial H(\beta,\theta)}{\partial \theta}\big),
\end{aligned}$$

where

$$\begin{aligned}
H(\beta,\theta) &= \int \Phi\big(\frac{\big[R_1(\theta,\mathbf{X},Y) - R_0(\beta,\mathbf{X})\big]}{\Sigma_R^{1/2}(\beta,\theta)}\big) dF_{1D}(\mathbf{X},Y) \\
\Sigma_R &= \big(\frac{\partial R_0(r,\mathbf{X})}{\partial r}|_{r=\beta}\big)^T \Sigma_0(\beta)\big(\frac{\partial R_0(r,\mathbf{X})}{\partial r}|_{r=\beta}\big) + \big(\frac{\partial R_1(s,\mathbf{X},Y)}{\partial s}|_{s=\theta}\big)^T \Sigma_1(\theta)\big(\frac{\partial R_1(s,\mathbf{X},Y)}{\partial s}|_{t=\theta}\big) \\
&\quad -2\big(\frac{\partial R_0(r,\mathbf{X})}{\partial r}|_{r=\beta}\big)^T cov\big(\sqrt{n}(\hat{\beta}-\beta), \sqrt{n}(\hat{\theta}-\theta)\big)\big(\frac{\partial R_1(s,\mathbf{X},Y)}{\partial s}|_{s=\theta}\big).
\end{aligned}$$

When $\hat{\beta}$ and $\hat{\theta}$ have influence function ψ_{0i} and ψ_{1i}, $\sqrt{n}(\hat{\beta}-\beta) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_{0i}(\mathbf{X}) + o_p(1)$, $\sqrt{n}(\hat{\theta}-\theta) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_{1i}(\mathbf{X},Y) + o_p(1)$, we have that,

$$\begin{aligned}
\Sigma_0(\beta) = var(\psi_{0i}(\mathbf{X})), \quad \Sigma_1(\theta) = var(\psi_{1i}(\mathbf{X},Y)) \\
cov\big(\sqrt{n}(\hat{\beta}-\beta), \sqrt{n}(\hat{\theta}-\theta)\big) = cov\big(\psi_{0i}(\mathbf{X}), \psi_{1i}(\mathbf{X},Y)\big).
\end{aligned}$$

3.4

Similarly, we can derive the asymptotic result for $\widehat{NRI^{>0}_{ne}}$.

$$\sqrt{n}\big[\widehat{NRI^{>0}_{ne}} - NRI^{>0}_{ne}\big] \rightarrow_d N(0, \Sigma_{NRI^{>0}_{ne}}),$$

$$\begin{aligned}
& \Sigma_{NRI^{>0}_{ne}} = var\big(\sqrt{n}\widehat{NRI^{>0}_{ne}}\big) \\
= \; & \frac{1-(NRI^{>0}_{ne})^2}{1-\rho} + 4\big(\frac{\partial G(\beta,\theta)}{\partial \beta}\big)^T \Sigma_0(\beta)\big(\frac{\partial G(\beta,\theta)}{\partial \beta}\big) + 4\big(\frac{\partial G(\beta,\theta)}{\partial \theta}\big)^T \Sigma_1(\theta)\big(\frac{\partial G(\beta,\theta)}{\partial \theta}\big) \\
& +8\big(\frac{\partial G(\beta,\theta)}{\partial \beta}\big)^T cov\big(\sqrt{n}(\hat{\beta}-\beta), \sqrt{n}(\hat{\theta}-\theta)\big)\big(\frac{\partial G(\beta,\theta)}{\partial \theta}\big),
\end{aligned}$$

where

$$\begin{aligned}
G(\beta,\theta) &= \int \Phi\big(\frac{\big[R_0(\beta,\mathbf{X}) - R_1(\theta,\mathbf{X},Y)\big]}{\Sigma_R^{1/2}(\beta,\theta)}\big) dF_{0\bar{D}}(\mathbf{X},Y) \\
\Sigma_R &= \big(\frac{\partial R_0(r,\mathbf{X})}{\partial r}|_{r=\beta}\big)^T \Sigma_0(\beta)\big(\frac{\partial R_0(r,\mathbf{X})}{\partial r}|_{r=\beta}\big) + \big(\frac{\partial R_1(s,\mathbf{X},Y)}{\partial s}|_{s=\theta}\big)^T \Sigma_1(\theta)\big(\frac{\partial R_1(s,\mathbf{X},Y)}{\partial s}|_{s=\theta}\big) \\
& \quad -2\big(\frac{\partial R_0(r,\mathbf{X})}{\partial r}|_{r=\beta}\big)^T cov\big(\sqrt{n}(\hat{\beta}-\beta), \sqrt{n}(\hat{\theta}-\theta)\big)\big(\frac{\partial R_1(s,\mathbf{X},Y)}{\partial s}|_{s=\theta}\big).
\end{aligned}$$

4 Asymptotic Distribution of a Two-Category NRI

In this section, we consider the two-category NRI calculated based on a reclassification table with risk threshold at t. NRI describes the "upward" and "downward" movements among risk categories comparing the new model to the old model. Let $\widehat{risk}_t(\mathbf{X}, Y)$ and $\widehat{risk}_t(\mathbf{X})$ denote the categorized risks with threshold t from the new model and from the old model,

then the two-category NRI^t is estimated as follows,

$$\begin{aligned}
\widehat{NRI_e^t} &= \hat{P}\big(\widehat{risk}_t(\mathbf{X},Y) > \widehat{risk}_t(\mathbf{X})|D=1\big) - \hat{P}\big(\widehat{risk}_t(\mathbf{X},Y) < \widehat{risk}_t(\mathbf{X})|D=1\big), \\
\widehat{NRI_{ne}^t} &= \hat{P}\big(\widehat{risk}_t(\mathbf{X},Y) < \widehat{risk}_t(\mathbf{X})|D=0\big) - \hat{P}\big(\widehat{risk}_t(\mathbf{X},Y) > \widehat{risk}_t(\mathbf{X})|D=0\big), \\
\widehat{NRI^t} &= \widehat{NRI_e^t} + \widehat{NRI_{ne}^t}.
\end{aligned}$$

With a single risk threshold at t, it can be shown that,

$$\begin{aligned}
NRI_e^t &= TPR(t, R_1(\theta, \mathbf{X}, Y)) - TPR(t, R_0(\beta, \mathbf{X})), \\
NRI_{ne}^t &= FPR(t, R_0(\beta, \mathbf{X})) - FPR(t, R_1(\theta, \mathbf{X}, Y)),
\end{aligned}$$

where TPR, FPR are the true positive rate and false positive rate.

4.1

We consider $\widehat{NRI_e^t}$ first.

$$
\begin{aligned}
& -\sqrt{n}\big(\widehat{NRI_e^t} - NRI_e^t\big) \\
= \; & -\sqrt{n}\big\{\widehat{TPR}(r, R_1) - \widehat{TPR}(r, R_0) - \big[TPR(t, R_1) - TPR(t, R_0)\big]\big\} \\
= \; & -\sqrt{n}\big\{TPR(t, R_1(\hat{\theta}, \mathbf{X}, Y)) - TPR(t, R_0(\hat{\beta}, \mathbf{X})) - \big[TPR(t, R_1(\theta, \mathbf{X}, Y)) - TPR(t, R_0(\beta, \mathbf{X}))\big]\big\} \\
= \; & -\sqrt{n}\big\{\big[1 - \hat{F}_{1D}(R_1^{-1}(\hat{\theta}, t))\big] - \big[1 - \hat{F}_{0D}(R_0^{-1}(\hat{\beta}, t))\big]\big\} \\
& +\sqrt{n}\big\{\big[1 - F_{1D}(R_1^{-1}(\theta, t))\big] - \big[1 - F_{0D}(R_0^{-1}(\beta, t))\big]\big\} \\
= \; & \sqrt{n}\big\{\big[\hat{F}_{1D}(R_1^{-1}(\hat{\theta}, t)) - \hat{F}_{0D}(R_0^{-1}(\hat{\beta}, t))\big] - \big[F_{1D}(R_1^{-1}(\theta, t)) - F_{0D}(R_0^{-1}(\beta, t))\big]\big\} \\
= \; & \sqrt{n}\big\{\big[\hat{F}_{1D}(R_1^{-1}(\theta, t)) - \hat{F}_{0D}(R_0^{-1}(\beta, t))\big] - \big[F_{1D}(R_1^{-1}(\theta, t)) - F_{0D}(R_0^{-1}(\beta, t))\big]\big\} \\
& +\sqrt{n}\big\{\big[F_{1D}(R_1^{-1}(\hat{\theta}, t)) - F_{0D}(R_0^{-1}(\hat{\beta}, t))\big] - \big[F_{1D}(R_1^{-1}(\theta, t)) - F_{0D}(R_0^{-1}(\beta, t))\big]\big\} + o_p(1) \\
& \text{(since } \sqrt{n}\big\{\big[\hat{F}_{1D}(R_1^{-1}(\hat{\theta}, t)) - \hat{F}_{0D}(R_0^{-1}(\hat{\beta}, t))\big] - \big[\hat{F}_{1D}(R_1^{-1}(\theta, t)) - \hat{F}_{0D}(R_0^{-1}(\beta, t))\big]\big\} \\
& -\sqrt{n}\big\{\big[F_{1D}(R_1^{-1}(\hat{\theta}, t)) - F_{0D}(R_0^{-1}(\hat{\beta}, t))\big] - \big[F_{1D}(R_1^{-1}(\theta, t)) - F_{0D}(R_0^{-1}(\beta, t))\big]\big\} \to_p 0, \\
& \text{due to the equicontinuity of the process and assumption 3.)} \\
= \; & C + D
\end{aligned}
$$

We know that $\sqrt{n}\big[\hat{F}_{1D}(t_1) - \hat{F}_{0D}(t_2) - \big(F_{1D}(t_1) - F_{0D}(t_2)\big)\big]$ is a Gaussian process, $\sqrt{n}\big[\hat{F}_{1D}(t_1) - \hat{F}_{0D}(t_2) - \big(F_{1D}(t_1) - F_{0D}(t_2)\big)\big] \to_d N(0, \big(F_{1D}(t_1) - F_{0D}(t_2)\big)\big[1 - \big(F_{1D}(t_1) - F_{0D}(t_2)\big)\big]/\rho)$, where $\rho = n_D/n$. Thus, the asymptotic variance of C is,

$$
\begin{aligned}
& var(C) \doteq -\frac{1}{\rho}\big[F_{1D}(R_1^{-1}(\theta, t)) - F_{0D}(R_0^{-1}(\beta, t))\big]\big\{1 + \big[F_{1D}(R_1^{-1}(\theta, t)) - F_{0D}(R_0^{-1}(\beta, t))\big]\big\} \\
= \; & \frac{1}{\rho}\big[TPR(t, R_1(\theta, \mathbf{X}, Y)) - TPR(t, R_0(\beta, \mathbf{X}))\big]\big\{1 - \big[TPR(t, R_1(\theta, \mathbf{X}, Y)) - TPR(t, R_0(\beta, \mathbf{X}))\big]\big\} \\
= \; & \frac{1}{\rho} NRI_e^t\big[1 - NRI_e^t\big]
\end{aligned}
$$

4.2

On the other hand, by Taylor expansion, we have,

$$\begin{aligned} D &= \sqrt{n}\big\{\big[F_{1D}(R_1^{-1}(\hat{\theta},t)) - F_{0D}(R_0^{-1}(\hat{\beta},t))\big] - \big[F_{1D}(R_1^{-1}(\theta,t)) - F_{0D}(R_0^{-1}(\beta,t))\big]\big\} \\ &= f_{1D}(R_1^{-1}(\theta,t))\big(\frac{\partial R_1^{-1}(\theta,t)}{\partial\theta}\big)^T\sqrt{n}(\hat{\theta}-\theta) - f_{0D}(R_0^{-1}(\beta,t))\big(\frac{\partial R_0^{-1}(\beta,t)}{\partial\beta}\big)^T\sqrt{n}(\hat{\beta}-\beta) + o_p(1) \\ &= \big(\frac{\partial TPR(t,R_0(\beta,\mathbf{X}))}{\partial\beta}\big)^T\sqrt{n}(\hat{\beta}-\beta) - \big(\frac{\partial TPR(t,R_1(\theta,\mathbf{X},Y))}{\partial\theta}\big)^T\sqrt{n}(\hat{\theta}-\theta) + o_p(1). \end{aligned}$$

Thus,

$$\begin{aligned} var(D) &= \big(\frac{\partial TPR(t,R_0(\beta,\mathbf{X}))}{\partial\beta}\big)^T\Sigma_0(\beta)\big(\frac{\partial TPR(t,R_0(\beta,\mathbf{X}))}{\partial\beta}\big) \\ &\quad + \big(\frac{\partial TPR(t,R_1(\theta,\mathbf{X},Y))}{\partial\theta}\big)^T\Sigma_1(\theta)\big(\frac{\partial TPR(t,R_1(\theta,\mathbf{X},Y))}{\partial\theta}\big) \\ &\quad - 2\big(\frac{\partial TPR(t,R_0(\beta,\mathbf{X}))}{\partial\beta}\big)^T cov\big(\sqrt{n}(\hat{\beta}-\beta),\sqrt{n}(\hat{\theta}-\theta)\big)\big(\frac{\partial TPR(t,R_1(\theta,\mathbf{X},Y))}{\partial\theta}\big). \end{aligned}$$

In the above, assumption 4 guarantees random variable D is not degenerate. In contrast to the derivation in section 3.2 of $NRI_e^{>0}$, this assumption is necessary for establishing asymptotic Normality of the two-category NRI. This is because when assumption 4 fails, $var(C) \doteq \frac{1}{\rho}NRI_e^t\big[1 - NRI_e^t\big] = 0$.

4.3

Next, we compute $cov(C,D)$. This includes covariances between $\sqrt{n}(\hat{\beta}-\beta)$, $\sqrt{n}(\hat{\theta}-\theta)$ and $\sqrt{n}(\hat{F}_{0D}-F_{0D})$, $\sqrt{n}(\hat{F}_{1D}-F_{1D})$. Denote the influence functions of $\hat{\beta}$ and $\hat{\theta}$ by ψ_{0i} and ψ_{1i}, respectively. We have $\sqrt{n}(\hat{\beta}-\beta) = \frac{1}{\sqrt{n}}\sum_{i=1}^n\psi_{0i}(\mathbf{X}) + o_p(1)$, $\sqrt{n}(\hat{\theta}-\theta) = \frac{1}{\sqrt{n}}\sum_{i=1}^n\psi_{1i}(\mathbf{X},Y) +$

$o_p(1)$. Thus,

$$
\begin{aligned}
& cov\big(\sqrt{n}(\hat{\beta}-\beta), \sqrt{n}(\hat{F}_{0D}(t)-F_{0D}(t))\big) \\
= \; & cov\big(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_{0i}(\mathbf{X}), \frac{1}{\sqrt{n}}\sum_{i=1}^{n} I(\mathbf{X}_D \le t) - F_{0D}(t)\big) \\
= \; & cov\big(\psi_{0i}(\mathbf{X}_D), I(\mathbf{X}_D \le t) - F_{0D}(t)\big) \quad \text{(since cases and controls are independent.)} \\
= \; & E\big(\psi_{0i}(\mathbf{X}_D)\big[I(\mathbf{X}_D \le t) - F_{0D}(t)\big]\big) \\
= \; & E\big(\psi_{0i}(\mathbf{X}_D)I(\mathbf{X}_D \le t)\big) - F_{0D}(t)E\big(\psi_{0i}(\mathbf{X}_D)\big) \equiv C_{\beta,0}(\mathbf{X}_D, t).
\end{aligned}
$$

Similarly, we can derive that,

$$
\begin{aligned}
& cov\big(\sqrt{n}(\hat{\beta}-\beta), \sqrt{n}(\hat{F}_{1D}(t)-F_{1D}(t))\big) \\
= \; & E\big(\psi_{0i}(\mathbf{X}_D)I(\mathbf{Z}_D \le t)\big) - F_{1D}(t)E\big(\psi_{0i}(\mathbf{X}_D)\big) \equiv C_{\beta,1}(\mathbf{Z}_D, t) = C_{\beta,1}(\mathbf{X}_D, Y_D, t), \\
& cov\big(\sqrt{n}(\hat{\theta}-\theta), \sqrt{n}(\hat{F}_{0D}(t)-F_{0D}(t))\big) \\
= \; & E\big(\psi_{1i}(\mathbf{Z}_D)I(\mathbf{X}_D \le t)\big) - F_{0D}(t)E\big(\psi_{1i}(\mathbf{Z}_D)\big) \equiv C_{\theta,0}(\mathbf{Z}_D, t) = C_{\theta,0}(\mathbf{X}_D, Y_D, t), \\
& cov\big(\sqrt{n}(\hat{\theta}-\theta), \sqrt{n}(\hat{F}_{1D}(t)-F_{1D}(t))\big) \\
= \; & E\big(\psi_{1i}(\mathbf{Z}_D)I(\mathbf{Z}_D \le t)\big) - F_{1D}(t)E\big(\psi_{1i}(\mathbf{Z}_D)\big) \equiv C_{\theta,1}(\mathbf{Z}_D, t) = C_{\theta,1}(\mathbf{X}_D, Y_D, t).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
& cov(C, D) \\
= \; & \big(\frac{\partial TPR(t, R_0(\beta, \mathbf{X}))}{\partial \beta}\big)^T cov\big(\sqrt{n}(\hat{\beta} - \beta), \sqrt{n}\big[\hat{F}_{1D}(R_1^{-1}(\theta, t)) - F_{1D}(R_1^{-1}(\theta, t))\big]\big) \\
& - \big(\frac{\partial TPR(t, R_0(\beta, \mathbf{X}))}{\partial \beta}\big)^T cov\big(\sqrt{n}(\hat{\beta} - \beta), \sqrt{n}\big[\hat{F}_{0D}(R_0^{-1}(\beta, t)) - F_{0D}(R_0^{-1}(\beta, t))\big]\big) \\
& - \big(\frac{\partial TPR(t, R_1(\theta, \mathbf{X}, Y))}{\partial \theta}\big)^T cov\big(\sqrt{n}(\hat{\theta} - \theta), \sqrt{n}\big[\hat{F}_{1D}(R_1^{-1}(\theta, t)) - F_{1D}(R_1^{-1}(\theta, t))\big]\big) \\
& + \big(\frac{\partial TPR(t, R_1(\theta, \mathbf{X}, Y))}{\partial \theta}\big)^T cov\big(\sqrt{n}(\hat{\theta} - \theta), \sqrt{n}\big[\hat{F}_{0D}(R_0^{-1}(\beta, t)) - F_{0D}(R_0^{-1}(\beta, t))\big]\big) \\
= \; & \big(\frac{\partial TPR(t, R_0)}{\partial \beta}\big)^T C_{\beta,1}(\mathbf{X}_D, Y_D, R_1^{-1}(\theta, t)) - \big(\frac{\partial TPR(t, R_0)}{\partial \beta}\big)^T C_{\beta,0}(\mathbf{X}_D, R_0^{-1}(\beta, t)) \\
& - \big(\frac{\partial TPR(t, R_1)}{\partial \theta}\big)^T C_{\theta,1}(\mathbf{X}_D, Y_D, R_1^{-1}(\theta, t)) + \big(\frac{\partial TPR(t, R_1)}{\partial \theta}\big)^T C_{\theta,0}(\mathbf{X}_D, Y_D, R_0^{-1}(\beta, t)).
\end{aligned}
$$

Since C and D are asymptotically Normal random variables, we have $\sqrt{n}\big(\widehat{NRI_e^t} - NRI_e^t\big)$, or equivalently $\sqrt{n}\big\{\widehat{TPR}(r, R_1) - \widehat{TPR}(r, R_0) - \big[TPR(t, R_1) - TPR(t, R_0)\big]\big\}$ is asymptotically Normal with mean 0 and variance $\Sigma_{NRI_e^t} = var(C) + var(D) + 2cov(C, D)$.

4.4

Similarly, we can derive asymptotic result for NRI_{ne}^t, or equivalently, $FPR(t, R_1(\theta, \mathbf{X}, Y)) - FPR(t, R_0(\beta, \mathbf{X}))$:

$$\sqrt{n}\big(\widehat{NRI_{ne}^t} - NRI_{ne}^t\big) \rightarrow_d N(0, \Sigma_{NRI_{ne}^t}),$$

where

$$\Sigma_{NRI^t_{ne}} = var(E) + var(F) + 2cov(E,F),$$

$$\begin{aligned}
E &= \sqrt{n}\big\{\big[\hat{F}_{1\bar{D}}(R_1^{-1}(\theta,t)) - \hat{F}_{0\bar{D}}(R_0^{-1}(\beta,t))\big] - \big[F_{1\bar{D}}(R_1^{-1}(\theta,t)) - F_{0\bar{D}}(R_0^{-1}(\beta,t))\big]\big\},\\
F &= \sqrt{n}\big\{\big[F_{1\bar{D}}(R_1^{-1}(\hat{\theta},t)) - F_{0\bar{D}}(R_0^{-1}(\hat{\beta},t))\big] - \big[F_{1\bar{D}}(R_1^{-1}(\theta,t)) - F_{0\bar{D}}(R_0^{-1}(\beta,t))\big]\big\}.
\end{aligned}$$

$$\begin{aligned}
var(E) &= \frac{1}{1-\rho} NRI^t_{ne}\big[1 - NRI^t_{ne}\big],\\
var(F) &= \Big(\frac{\partial FPR(t, R_0(\beta,\mathbf{X}))}{\partial\beta}\Big)^T \Sigma_0(\beta) \Big(\frac{\partial FPR(t, R_0(\beta,\mathbf{X}))}{\partial\beta}\Big)\\
&\quad + \Big(\frac{\partial FPR(t, R_1(\theta,\mathbf{X},Y))}{\partial\theta}\Big)^T \Sigma_1(\theta) \Big(\frac{\partial FPR(t, R_1(\theta,\mathbf{X},Y))}{\partial\theta}\Big)\\
&\quad - 2\Big(\frac{\partial FPR(t, R_0(\beta,\mathbf{X}))}{\partial\beta}\Big)^T cov\big(\sqrt{n}(\hat{\beta}-\beta), \sqrt{n}(\hat{\theta}-\theta)\big) \Big(\frac{\partial FPR(t, R_1(\theta,\mathbf{X},Y))}{\partial\theta}\Big),\\
cov(E,F) &= \Big(\frac{\partial FPR(t, R_0)}{\partial\beta}\Big)^T \bar{C}_{\beta,1}(\mathbf{X}_D, Y_D, R_1^{-1}(\theta,t)) - \Big(\frac{\partial FPR(t, R_0)}{\partial\beta}\Big)^T \bar{C}_{\beta,0}(\mathbf{X}_D, R_0^{-1}(\beta,t))\\
&\quad - \Big(\frac{\partial FPR(t, R_1)}{\partial\theta}\Big)^T \bar{C}_{\theta,1}(\mathbf{X}_D, Y_D, R_1^{-1}(\theta,t)) + \Big(\frac{\partial FPR(t, R_1)}{\partial\theta}\Big)^T \bar{C}_{\theta,0}(\mathbf{X}_D, Y_D, R_0^{-1}(\beta,t)).
\end{aligned}$$

$$\begin{aligned}
\bar{C}_{\beta,0}(\mathbf{X}_{\bar{D}}, t) &= E\big(\psi_{0i}(\mathbf{X}_{\bar{D}}) I(\mathbf{X}_{\bar{D}} \le t)\big) - F_{0\bar{D}}(t) E\big(\psi_{0i}(\mathbf{X}_{\bar{D}})\big),\\
\bar{C}_{\beta,1}(\mathbf{X}_{\bar{D}}, Y_{\bar{D}}, t) &= E\big(\psi_{0i}(\mathbf{X}_{\bar{D}}) I(\mathbf{Z}_{\bar{D}} \le t)\big) - F_{1\bar{D}}(t) E\big(\psi_{0i}(\mathbf{X}_{\bar{D}})\big),\\
\bar{C}_{\theta,0}(\mathbf{X}_{\bar{D}}, Y_{\bar{D}}, t) &= E\big(\psi_{1i}(\mathbf{Z}_{\bar{D}}) I(\mathbf{X}_{\bar{D}} \le t)\big) - F_{0\bar{D}}(t) E\big(\psi_{1i}(\mathbf{Z}_{\bar{D}})\big),\\
\bar{C}_{\theta,1}(\mathbf{X}_{\bar{D}}, Y_{\bar{D}}, t) &= E\big(\psi_{1i}(\mathbf{Z}_{\bar{D}}) I(\mathbf{Z}_{\bar{D}} \le t)\big) - F_{1\bar{D}}(t) E\big(\psi_{1i}(\mathbf{Z}_{\bar{D}})\big).
\end{aligned}$$

5 Sampling Distributions in Finite Samples

We have established asymptotic Normality for category-free and two-category NRI_e and NRI_{ne}. Another important question is how accurate the Normal approximation is in finite samples. In this section, we perform simulations to examine sampling distributions of NRI

statistics with quantile-quantile (QQ) plots.

We simulate data with risks following a logistic model since logistic regression is commonly used in risk modeling. The same family of simulation models have been used by several researchers to examine methodology related to incremental value [3, 4, 20]. Specifically, Let ρ denote the disease prevalence and n denote the sample size. We first generate $D \sim Binomial(N, \rho)$. We then generate the baseline marker $X \sim N(0,1)$ and the new marker $Y \sim N(0,1)$, independently in controls, and $X \sim N(\mu_X, 1)$, $Y \sim N(\mu_Y, 1)$, independently in cases. This guarantees the logistic model holds when X alone or X and Y are included as predictors. This can be seen by Bayes rule,

$$\begin{aligned} P(D=1|X=x) &= \frac{P(X=x|D=1)P(D=1)}{P(X=x|D=1)P(D=1)+P(X=x|D=0)P(D=0)} \\ &= \frac{\rho e^{-(x-\mu_X)^2/2}}{\rho e^{-(x-\mu_X)^2/2}+(1-\rho)e^{-x^2/2}}, \\ logitP(D=1|X=x) &= log\frac{\rho}{1-\rho}e^{x^2/2-(x-\mu_X)^2/2} \\ &= \mu_X x - \frac{\mu_X^2}{2} + log\frac{\rho}{1-\rho}. \end{aligned}$$

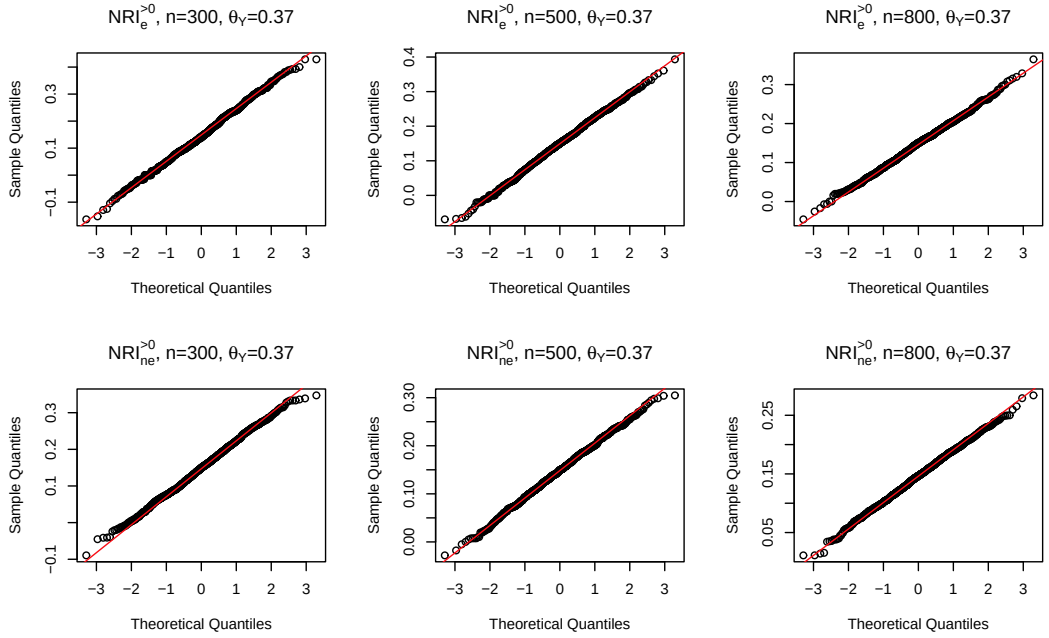
Similarly we have,

$$logitP(D=1|X=x,Y=y) = \mu_X x + \mu_Y y - \frac{\mu_X^2+\mu_Y^2}{2} + log\frac{\rho}{1-\rho}.$$

Thus the logistic model holds. Moreover, the coefficients in the models are $\beta_X = \theta_X = \mu_X$ and $\theta_Y = \mu_Y$.

For our simulations, we always set $\rho = 0.2$ and $\mu_X = 0.74$, while $\mu_Y = 0, 0.37$ or 0.74 to reflect a new marker with no, modest or relatively large predictive strength. We consider both the category-free NRI and the two-category NRI. For the two-category NRI, we used the 20% quantile of $R_0(\beta, X)$ as risk threshold, which means the proportion of subjects classified as "high risk" by the baseline risk model is about the same as the prevalence. This

Figure 1: QQ plots of $NRI_e^{>0}$ and $NRI_{ne}^{>0}$ for a new marker with a modest incremental value under various sample size.

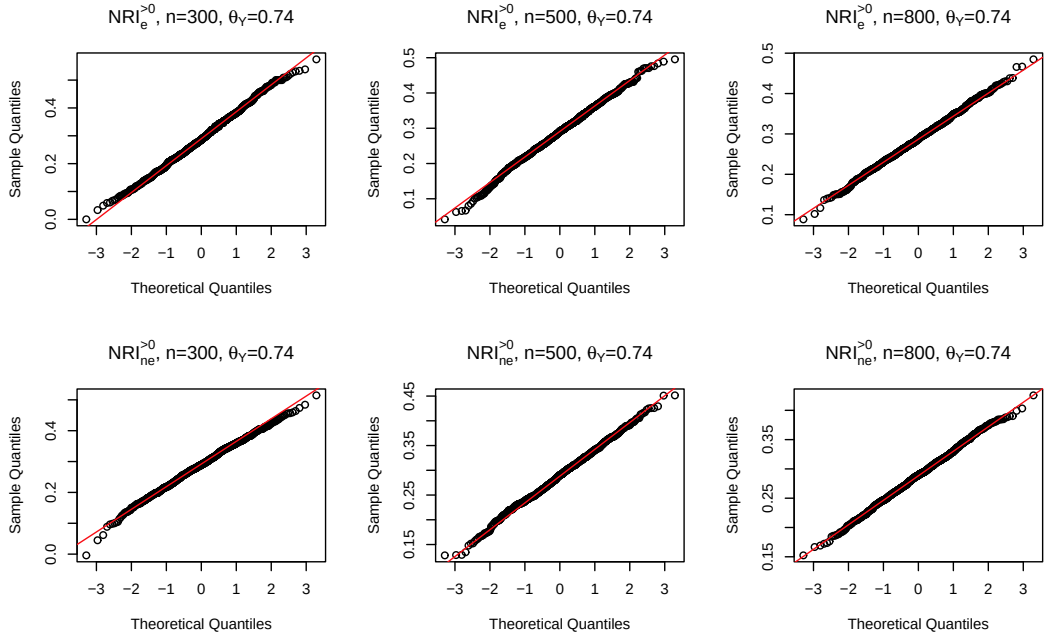


threshold was obtained empirically by simulating a data set of size 10,000,000.

Figure 1 is QQ plots of $NRI_e^{>0}$ and $NRI_{ne}^{>0}$ for a new marker with a modest incremental value under various sample size. Figure 2 is QQ plots of $NRI_e^{>0}$ and $NRI_{ne}^{>0}$ when the new marker has a relatively large incremental value. In both figures, the upper panel is for $NRI_e^{>0}$ and the bottom panel is for $NRI_{ne}^{>0}$. Sample size increases from left to right. In both situations (modest or large incremental value), sampling distributions of $NRI_e^{>0}$ and $NRI_{ne}^{>0}$ are quite close to Normal with a relatively small sample size, and the approximation improves as sample size increases.

Figure 3 and Figure 4 are QQ plots of the two-category NRI with threshold at 20% risk quantile for a new marker with a modest incremental value or with a relatively large incremental value. Because of the discreteness of the two-category NRI introduced by the risk threshold, a larger sample size is expected for asymptotic theory to take effect compared to

Figure 2: QQ plots of $NRI_e^{>0}$ and $NRI_{ne}^{>0}$ for a new marker with a relatively large incremental value under various sample size.



the category-free NRI. In addition, note that the sample sizes of Figure 3 are 1,000, 3,000 and 5,000, larger than the sample sizes of Figure 4, which are 1,000, 2,000 and 3,000. We chose these sample sizes so that the plots can better characterize the distributional behavior of $NRI_e^{0.2}$ and $NRI_{ne}^{0.2}$ as they move towards their limiting distributions. In both figures, we can see the discreteness in the QQ plots. For the same sample size, sampling distribution of the two-category NRI is closer to Normal when the new marker has a relatively large incremental value than when it has only a modest incremental value. For example, at sample size 3,000, the distributions of $NRI_e^{0.2}$ and $NRI_{ne}^{0.2}$ are relatively close to Normal in Figure 4 when the new marker has a relatively large incremental value, while in Figure 3, when the incremental value is modest, the distributions of $NRI_e^{0.2}$ and $NRI_{ne}^{0.2}$ still have visible deviation from Normal. As one expects, the Normal approximation becomes better as sample size increases. However, compared with the category-free NRI, a much larger sample size is required for sampling distributions of $NRI_e^{0.2}$ and $NRI_{ne}^{0.2}$ be to approximately Normal.

Figure 3: QQ plots of $NRI_e^{0.2}$ and $NRI_{ne}^{0.2}$ for a new marker with a modest incremental value under various sample size.

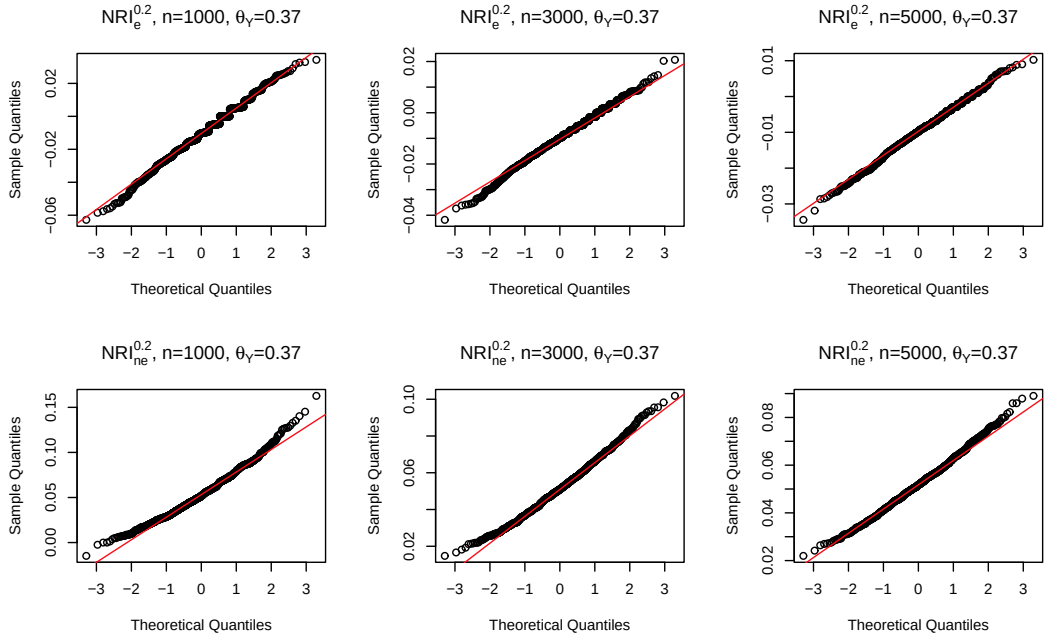


Figure 4: QQ plots of $NRI_e^{0.2}$ and $NRI_{ne}^{0.2}$ for a new marker with a relatively large incremental value under various sample size.

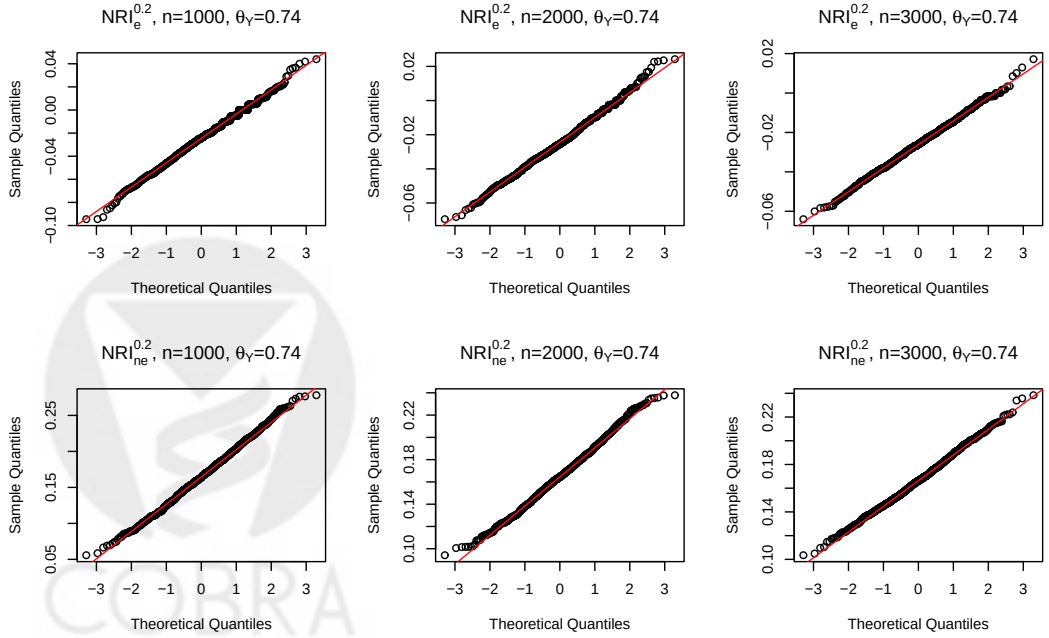
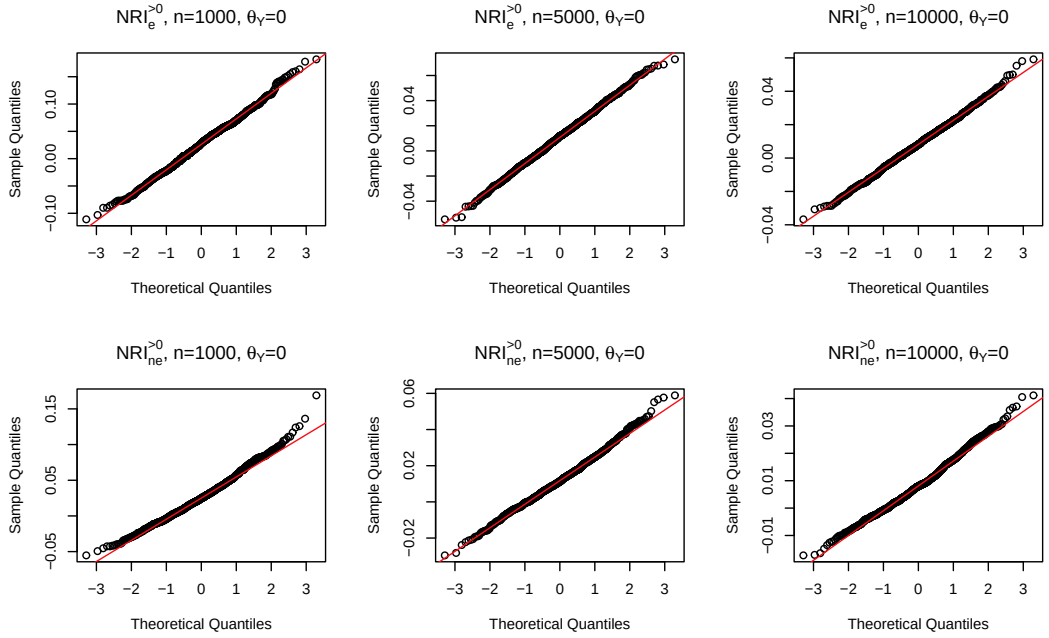


Figure 5: QQ plots of $NRI_e^{>0}$ and $NRI_{ne}^{>0}$ for a new marker with no incremental value under various sample size.



An unexpected finding suggested by Figure 3 is that, although the number of cases is much smaller than the number of controls in our simulation settings, sampling distribution of $NRI_{ne}^{0.2}$ does not approach Normality faster than that of $NRI_e^{0.2}$. Their distributional behavior appears to depend more on the total sample size than on the numbers of cases or controls.

As noted in section 3.2, assumption 4 is not essential for the sampling distribution of category-free NRI to be asymptotic Normal. Here, we also examine the finite sample distributional behavior of $NRI_e^{>0}$ and $NRI_{ne}^{>0}$ when the new marker has no incremental value at all. The QQ plots are shown in Figure 5. The results suggest that, when the new marker has an incremental value on the boundary zero, the category-free NRI statistics still have asymptotically Normality, but a much larger sample size is needed to reach a good Normal approximation compared with situations where the incremental value is away from zero.

6 Discussion

In this paper, we examined the asymptotic and finite sample distributional behavior of NRI statistics when risks are estimated from risk models fit to a dataset. We established asymptotic Normality of NRI statistics, which provides some justification for constructing confidence intervals via Normal approximation. For the category-free NRI, asymptotic Normality can be reached with a rather small sample size when the new marker has an incremental value away from zero. However, a large sample size is required for the two-category event or nonevent NRI to get close to Normal, especially when the new marker has only a modest incremental value. Moreover, a nonzero value of the incremental value of the new marker (assumption 4) is a necessary assumption for establishing asymptotic Normality for the two-category NRI. When this condition fails, simulation results suggest that the limiting distribution of NRI_e^t and NRI_{ne}^t is not Normal (results not shown). This is similar to some other measures of incremental value, such as the integrated discrimination improvement index [21]. For the category-free NRI, although this assumption is not necessary for establishing its asymptotic Normality, a much larger sample size is needed for the distribution to be approximately Normal compared with situations when the incremental value is away from zero.

However, the behavior of NRI statistics on the boundary zero, i.e., when the new marker has no incremental value, is not the focus of our paper. This is because quantifying predictive improvement is more of interests when the new marker is useful. Otherwise, for a marker with unknown predictive ability, it is advised to first perform a test to determine whether this marker has nonzero incremental value [3]. Nevertheless, we do not need to construct such a test based on $NRI^{>0}$ statistics, because the null hypothesis of zero value of $NRI^{>0}$ is equivalent to the null hypothesis that the coefficient of the marker in the risk model is zero, for which more powerful tests can be constructed [3].

In the derivation, we assumed the risk models used reflect the truth and did not consider model misspecification. However, as argued by Pepe et al [3], poorly calibrated models are not acceptable for risk prediction. The performance characteristics of a risk model should be examined first. Only adequately calibrated models should advance to further evaluation and potential adoption.

Acknowledgements

The author thanks Dr. Kathleen Kerr and Dr. Margaret Pepe for their valuable advice and their insightful discussion.

References

1. Pencina MJ, D'Agostino RB, Sr, D'Agostino RB, Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Statistics in Medicine* 2008; **27**:157-172, DOI: 10.1002/sim.2929.

2. Pencina MJ, D'Agostino RB, Sr, Steyerberg EW. Extensions of net reclassi?cation improvement calculations to measure usefulness of new biomarkers. *Statistics in Medicine* 2010; **30**:11-21, DOI: 10.1002/sim.4085.

3. Pepe MS, Kerr KF, Longton G, and Wang Z. Testing for improvement in prediction model performance. *Statistics in Medicine*, (in press), 2013. 4. Kerr KF, Wang Z, Janes H, McClelland RL, Psaty BM, Wu C and Pepe MS. Measuring the prediction increment with net reclassification indices. 2013, *submitted*.

5. Barlow WE, White E, Ballard-Barbash R, Vacek PM, Titus-Ernstoff L, Carney PA, Tice JA, Buist DS, Geller BM, Rosenberg R, Yankaskas BC, Kerlikowske K. Prospective breast cancer risk prediction model for women undergoing screening mammography.

Journal of the National Cancer Institute 2006; **98**:1204-1214, DOI: 10.1093/jnci/djj331.

6. Chen J, Pee D, Ayyagari R, Graubard B, Schairer C, Byrne C, Benichou J, Gail MH. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. *Journal of the National Cancer Institute* 2006; **98**:1215-1226, DOI: 10.1093/jnci/djj332.

7. Wacholder S, Hartge P, Prentice R, Garcia-Closas M, Feigelson HS, Diver WR, Thun MJ, Cox DG, Hankinson SE, Kraft P, Rosner B, Berg CD, Brinton LA, Lissowska J, Sherman ME, Chlebowski R, Kooperberg C, Jackson RD, Buckman DW, Hui P, Pfeiffer R, Jacobs KB, Thomas GD, Hoover RN, Gail MH, Chanock SJ, Hunter DJ. Performance of common genetic variants in breast-cancer risk models. *New England Journal of Medicine* 2010; **362**:986-993, 10.1056/NEJMoa0907727.

8. Gail MH. Probability discriminatory accuracy from single-nucleotide polymorphisms in models to predict breast cancer risk. *Journal of the National Cancer Institute* 2008; **100**:1037-41,DOI: 10.1093/jnci/djn180.

9. Gail MH. Value of adding single-nucleotide polymorphism genotypes to a breast cancer risk model. *Journal of the National Cancer Institute* 2009; **101**:959-963, DOI: 10.1093/jnci/djp130

10. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, Mulvihill JJ. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *Journal of the National Cancer Institute* 1989; **81**:1879-1886, DOI: 10.1093/jnci/81.24.1879.

11. Gail MH, Costantino JP. Validating and improving models for projecting the absolute risk of breast cancer. *Journal of the National Cancer Institute* 2001; **93**:334-335, DOI: 10.1093/jnci/93.5.334

12. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998; **97**:1837-1847, DOI: 10.1161/01.CIR.97.18.1837.

13. Gu W, Pepe, MS. Measures to summarize and compare the predictive capacity of markers. *The International Journal of Biostatistics* 2009; **5**: article 27.

14. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007;**115**(7): 928-935.

15. Pepe MS, Janes HE. Gauging the performance of SNPs, biomarkers, and clinical factors for predicting risk of breast cancer. *Journal of the National Cancer Institute*. 2008;**100**(14):978-979.

16. Pepe MS. Problems with risk reclassification methods for evaluating prediction models. *American Journal of Epidemiology*. 2011; **173**(11): 1327-1335.

17. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, and Kattan MW. Assessing the performance of prediction models: A framework for traditional and novel measures. *Epidemiology*. 2010; **21**(1): 128-38.

18. Cook NR, Ridker PM. Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassification measures. *Annals of internal medicine*, 2009; **150**(11):795-802.

19. Polonsky TS, McClelland RL, Jorgensen NW, and et al. Coronary artery calcium score and risk classification for coronary heart disease prediction. *JAMA: The Journal of the American Medical Association*, 2010; **303**(16):1610-1616, doi: 10.1001/jama.2010.461.

20. Pencina MJ, D'Agostino RB, Pencina KM, Janssens AC, Greenland P. Interpreting incremental value of markers added to risk prediction models. *American Journal of Epidemiology*, 2012; **176**(6):473-481.

21. Kerr KF, McClelland RL, Brown ER, Lumley T. Evaluating the incremental value of new biomarkers with integrated discrimination improvement. *American Journal of Epidemiology*, 2011; **174**(3):364-374.