Published in final edited form as:

Stat Med. 2009 June 15; 28(13): 1841–1860. doi:10.1002/sim.3575.

Generalized ROC curve inference for a biomarker subject to a limit of detection and measurement error†

Neil J. Perkins^{1,*†}, Enrique F. Schisterman¹, and Albert Vexler^{1,2}

¹Division of Epidemiology, Statistics and Prevention Research, National Institute of Child Health and Human Development, National Institutes of Health, DHHS, 6100 Executive BLVD, Rm 7B03, Bethesda, MD 20852, U.S.A.

²Department of Biostatistics, State University of New York at Buffalo, Buffalo, NY 14214, U.S.A.

SUMMARY

The receiver operating characteristic (ROC) curve is a tool commonly used to evaluate biomarker utility in clinical diagnosis of disease, especially during biomarker development research. Emerging biomarkers are often measured with random measurement error and subject to limits of detection that hinder their potential utility or mask an ability to discriminate by negatively biasing the estimates of ROC curves and subsequent area under the curve. Methods have been developed to correct the ROC curve for each of these types of sources of bias but here we develop a method by which the ROC curve is corrected for both simultaneously through replicate measures and maximum likelihood. Our method is evaluated via simulation study and applied to two potential discriminators of women with and without preeclampsia.

Keywords

ROC curve; limit of detection; measurement error; area under the curve; replicates

1. INTRODUCTION

Diagnosis and prognosis of diseases are increasingly dependent on the effectiveness biological markers to discriminate individuals with and without disease. The receiver operating characteristic (ROC) curve is a statistical device commonly used to evaluate the utility of a potential biomarker.

Consider populations of individuals with and without a particular disease. Let independent random variables X and Y , respectively, denote the levels of a specific biomarker of individuals within these populations which are governed by cumulative distribution functions $F(x)$ and $G(y)$. Suppose, the biomarker is thought to have utility as an indicator of disease status with levels above some cut point, c , indicating individuals testing positive for the disease and levels below c corresponding those individuals testing negative. Two quantities useful in assessing biomarker utility, sensitivity ($q(c)$) and specificity ($p(c)$), are then $1-F(c)$ and $G(c)$, respectively, for a given c . The ROC curve is a mapping of $(p(c), q(c))$ across all possible c . Proposed uses for the ROC curve include assessing overall discriminatory ability, or over a

†This article is a U.S. Government work and is in the public domain in the U.S.A.

*Correspondence to: Neil J. Perkins, Division of Epidemiology, Statistics and Prevention Research, National Institute of Child Health and Human Development, National Institutes of Health, DHHS, 6100 Executive BLVD, Rm 7B03, Bethesda, MD 20852, U.S.A..

†perkinsn@mail.nih.gov.

specific range of $q(c)$ or $p(c)$ and the maximum ability to differentiate between the populations [1,2]. The area under the ROC curve, denoted here by AUC, is the most commonly used summary measure and is equal to $P(X > Y)$ [1,2]. The AUC measures separation between diseased and non-diseased populations based on biomarker levels. Ranging from 0.5 to 1, larger AUC values indicate greater separation and thus greater utility as a diagnostic tool.

A major complication in the development of emerging biomarkers is an investigator's ability to adequately assess biomarker utility from data with measurement error (ME). Conventional ROC methods assume that the biomarker levels measured are the true levels of those individuals. However, most biomarkers are subject to random additive ME or random variability from this true level of an individual. This variation could be due to intra-individual or measurement variability among other causes. Regardless of the origin, biomarker levels are often measured with error what is commonly assumed to be mean zero normally distributed error. Subsequently, measurements may be subject a limit of detection (LOD) where due to instrumental limitations measurements are undetectable below some level for that particular biomarker.

Formally, in the presence of ME the 'true' value of the biomarker, X and Y , is not available. Instead we assume that actual observations are the true biomarker levels plus some additional error, $\varepsilon_{X_i}, \varepsilon_{Y_j} \sim N(0, \sigma_\varepsilon^2)$,

$$\begin{aligned} x_i &= X_i + \varepsilon_{X,i}, & i=1, \dots, n_x \\ y_j &= Y_j + \varepsilon_{Y,j}, & j=1, \dots, n_y \end{aligned}$$

with X_i , Y_j , $\varepsilon_{X,i}$ and $\varepsilon_{Y,j}$ independent of each other. The equal variance assumption of $\varepsilon_{X,i}$ and $\varepsilon_{Y,j}$ is not necessary, but in many instances, justified because the variations leading to ME are due to factors connected with the laboratory or measurement process and do not depend on the individual's risk or disease status or the true biomarker value. When the true levels are normally distributed, adding independent normally distributed error results in levels that are also normally distributed with mean μ_X or μ_Y and variance $\sigma_X^2 + \sigma_\varepsilon^2$ or $\sigma_Y^2 + \sigma_\varepsilon^2$, respectively, depending on disease status.

Naively assuming that these are the true biomarker levels would result in an ROC curve and associated AUC estimate that is biased toward the null [3]. These negatively biased results hinder an investigator's ability to identify a biomarker as potentially useful and might lead to it being mislabeled useless for disease diagnosis. Instead, replicate measurements x_{ik} and y_{jk} where $k = 1, \dots, p$, are commonly used in order to correct for ME by estimating the amount of intra-individual variability present [3,4].

Estimation is often further complicated by the inability of laboratory methods or machinery to detect biomarker levels below some LOD, quantified as d , effectively censoring the data below d . That is to say that we observe

$$z_{x,i} = \begin{cases} X_i + \varepsilon_{X,i} & \text{if } X_i + \varepsilon_{X,i} \geq d \\ N/A & \text{if } X_i + \varepsilon_{X,i} < d \end{cases} \quad \text{and} \quad z_{y,j} = \begin{cases} Y_j + \varepsilon_{Y,j} & \text{if } Y_j + \varepsilon_{Y,j} \geq d \\ N/A & \text{if } Y_j + \varepsilon_{Y,j} < d \end{cases}$$

Obviously, omitting these unquantifiable values yields a bias assessment of biomarker effectiveness.

Substituting a replacement value, 0, $d/2$, $d/\sqrt{2}$ and d for unobservable data has been suggested as an alternative to provide a less biased assessment than omission [5]. It has been shown that all of these methods lead to biased estimation of mean and standard deviation parameters, regression coefficients, odds ratio, and AUC for a single biomarker [6–9]. Maximum likelihood estimators (MLE) of ROC measures have been developed based on univariate and multivariate likelihoods functions for these censored data sets and have shown to be efficient and asymptotically unbiased [8–12]. However, none of these methods have considered estimation based on data affected by both.

In this paper, we assume the availability of normally distributed biomarker levels, in replicate from populations with and without disease, affected by ME and LOD. Based on these data, we propose in Section 2 to generate MLEs for μ_X , σ_X^2 , μ_Y , σ_Y^2 and σ_ε^2 from a likelihood function that blends the general form of Lynn's bivariate normal developments for parameter estimates based on a sample with left censoring and Schisterman *et al.*, Resier and Faraggi's likelihood for replicates of normal biomarker levels with ME [3,4,13]. Subsequently, an MLE AUC will be generated based on MLEs of normal parameter that governs the distributions of true biomarker levels. In Section 3, the asymptotic distribution of AUC is considered and a confidence interval (CI) is developed. Point and CI estimators are evaluated via simulation in Section 4. An example in Section 5 using polychlorinated biphenyl levels to classify women with endometriosis is used to illustrate the impact of replacement and maximum likelihood techniques to correct for ME and LOD. We conclude with a brief discussion of issues surrounding estimation based on data affected by ME and LOD.

2. METHODS

Methods that correct for ME due so through some form of repeated measurements in order to assess σ_ε^2 . Here, correcting for ME is only half the battle but we will continue to assume the availability of repeated measurements on every individual in our data set. Subsequently, we can form estimators corrected for both sources of error using maximum likelihood.

Consider $p = 2$ repeated measurements for every individual in a study population of individuals with and without disease. Let $x_{ik} = X_i + \varepsilon_{X,ik}$ and $y_{jk} = Y_j + \varepsilon_{Y,jk}$ denote the k th, $k = 1$ or 2 , measurement of the i th and j th individuals, respectively, and $\mathbf{x}_i = (x_{i1}, x_{i2})$ and $\mathbf{y}_j = (y_{j1}, y_{j2})$ are bivariate vectors of observations. Using replicates in order to correct for ME was considered separately by Reiser and Faraggi and later by Schisterman [3,4,13]. The correlated replicates of the j th non-diseased individual follow the bivariate distribution

$$\mathbf{y}_j \sim N_2(\mu_Y, \Sigma_Y), \quad \text{where } \mu_Y = (\mu_Y, \mu_Y) \quad \text{and} \quad \Sigma_Y = \begin{bmatrix} \sigma_Y^2 + \sigma_\varepsilon^2 & \sigma_Y^2 \\ \sigma_Y^2 & \sigma_Y^2 + \sigma_\varepsilon^2 \end{bmatrix}$$

with correlation

$$\rho = \frac{\sigma_Y^2}{\sigma_Y^2 + \sigma_\varepsilon^2}$$

Replicates of biomarker levels in diseased individuals are similarly distributed. MLE's are generated by maximizing the standard likelihood function for binormal data, accounting for the singular mean and specific covariance structure.

Now let all measurement be subject to an LOD. Vectors $\vec{z}_{xi} = (z_{xi1}, z_{xi2})$ and $\vec{z}_{yi} = (z_{yi1}, z_{yi2})$ are observed replications on diseased and health persons, respectively, after the transformation in Section 1. The MLE's developed previously do not account for this additional source of error.

A maximum likelihood method does exist for multiple correlated normally distributed biomarkers affected by LODs in order to generate estimates of the parameters $\mu_1, \mu_2, \sigma_1, \sigma_2$ and ρ [8,14]. As the replicates are bivariate normally distributed, we can correct for the ME and LOD here by adapting this general likelihood by letting $\mu_1 = \mu_Y, \mu_2 = \mu_Y$,

$\sigma_1 = \sqrt{\sigma_Y^2 + \sigma_\varepsilon^2}, \sigma_2 = \sqrt{\sigma_Y^2 + \sigma_\varepsilon^2}$ and again, $\rho = \sigma_Y^2 / (\sigma_Y^2 + \sigma_\varepsilon^2)$. Given the two by n_Y matrix, $\mathbf{Z}_Y$, the general likelihood in the presence of ME and LOD is

$$L(\vec{\mu}_Y, \Sigma_Y; \mathbf{Z}_Y) = \prod_{j=1}^{n_Y} \left[\frac{1}{\sigma_\varepsilon \sqrt{\sigma_\varepsilon^2 + 2\sigma_Y^2}} \phi(\eta_{Y1j}) \phi(\eta_{Y2|Y1j}^*) \right]^{I(Z_{Y1j}, Z_{Y2j} \neq N/A)} \\ \times \left[\frac{1}{\sqrt{\sigma_\varepsilon^2 + \sigma_Y^2}} \phi(\eta_{Y1j}) \Phi(\eta_{d2|Y1j}^*) \right]^{I(Z_{Y1j} \neq N/A, Z_{Y2j} = N/A)} \\ \times \left[\frac{1}{\sqrt{\sigma_\varepsilon^2 + \sigma_Y^2}} \phi(\eta_{Y2j}) \Phi(\eta_{d1|Y2j}^*) \right]^{I(Z_{Y1j} = N/A, Z_{Y2j} \neq N/A)} \\ \times \left[\int_{-\infty}^{d_1} \frac{1}{\sqrt{\sigma_\varepsilon^2 + \sigma_Y^2}} \phi(\eta_{Y1j}) \Phi(\eta_{d2|Y1j}^*) dy_1 \right]^{I(Z_{Y1j}, Z_{Y2j} = N/A)} \quad (1)$$

where

$$\eta_{Y1j} = \frac{y_{j1} - \mu_Y}{\sqrt{\sigma_Y^2 + \sigma_\varepsilon^2}}$$

and

$$\eta_{Y2|Y1j}^* = \frac{\eta_{Y2j} - \rho \eta_{Y1j}}{\sqrt{1 - \rho^2}}$$

for the non-diseased parameters.

Maximizing the log of Equation (1) for both the diseased and non-diseased biomarker levels yields MLE's $\hat{\mu}_x, \hat{\mu}_y, \hat{\sigma}_x^2, \hat{\sigma}_y^2$ and $\hat{\sigma}_\varepsilon^2$. Using these MLE's, we can construct the MLE

$$A\hat{U}C = \Phi \left(\frac{\hat{\mu}_x - \hat{\mu}_y}{\sqrt{\hat{\sigma}_x^2 + \hat{\sigma}_y^2}} \right)$$

that is corrected for both ME and an LOD, that is to say $A\hat{U}C$ estimates the biomarkers true underlying ability to discriminate individuals with and without the disease.

3. ESTIMATOR DISTRIBUTION

Similar to previous asymptotic developments, the MLE's $\widehat{\mu}_X$, $\widehat{\mu}_Y$, $\widehat{\sigma}_X^2$, $\widehat{\sigma}_Y^2$, $\widehat{\sigma}_\varepsilon^2$, and $\widehat{AUC}$ constructed here are asymptotically normally distributed and unbiased [9,15]. As ultimately the AUC is our parameter of interest we will focus on the distribution of $\widehat{AUC}$, specifically $\sqrt{N}(\widehat{AUC} - AUC) \sim \text{Normal}(0, \sigma_A^2)$, where $N = n_X + n_Y$ and σ_A^2 is the variance parameter. The development of σ_A^2 here will mirror those in the appendix of Perkins *et al.* differing only in that Fisher's information matrix, denoted by $\mathbf{I}$, for $\widehat{\mu}_X$, $\widehat{\sigma}_X^2$ and $\widehat{\mu}_Y$, $\widehat{\sigma}_Y^2$, pair wise [9]. The derivatives and subsequent expected values necessary for $\mathbf{I}$ here can be achieved by accounting for the specific covariate structure in those developed for the general lower censored bivariate normal [15]. The details of these developments are left to the appendix.

Utilizing these distributional properties, α level two-tailed CI for AUC is constructed from

$$\widehat{AUC} \pm z_{\alpha/2} \frac{\sqrt{\sigma_A^2}}{\sqrt{N}}$$

As is often the case, when σ_A^2 is unknown, an approximate α level CI for AUC is formed by using the estimator $\widehat{\sigma}_A^2$ in lieu of σ_A^2 .

4. EVALUATION

A simulation study assesses the estimators developed in Sections 2 and 3, MLE for AUC that corrects for both an LOD and random additive ME. The $\widehat{AUC}$ was assessed using bias and root mean-squared error, RMSE, over $B = 2000$ iterations. Coverage probabilities of 95 per cent CIs for AUC's were recorded and examined as well. Sample sizes of varying magnitude, $n_X = n_Y = 50, 100, 1000$, were used to show the asymptotic nature of these estimators.

Simulated true non-diseased values were generated from a standard normal distribution, $\mu_Y = 0$ and $\sigma_Y = 1$. True diseased values were generated from a normal distribution with parameters $\sigma_X^2 = 1.0$ or 0.5 , equal and unequal variance cases, respectively, and mean μ_X corresponding to $AUC = 0.6, 0.7, 0.8, 0.9$.

ME was then introduced from independent mean zero normally distributions with variance σ_ε^2 . Error variability was linked to inherent biomarker variability, σ_Y^2 and σ_X^2 , through the 'reliability index' extended to two populations, [16]

$$R = \frac{\sigma_X^2 + \sigma_Y^2}{\sigma_X^2 + \sigma_Y^2 + 2\sigma_\varepsilon^2}$$

R is an index commonly used in the social science literature to reflect the proportion of observed variability that is due to the actual population [17]. Of course, $1 - R$ is the amount of observed variability that is due to error. Several scenarios of increasing error variability were considered $R = 0.90, 0.75, 0.50, 0.25$. As the method in Section 3.2 also corrects for an LOD, we set d based on the desired per cent, 0, 10, 25, 50, 75 per cent, of the non-diseased population, observed with ME, that would be unquantifiable, missing.

Tables I–VI show the results of our simulations. Bias of the equal and unequal variances cases are displayed in Tables I, II for every combination of sample size, true AUC, R and LOD. Tables III, IV and V, VI are for the same parameters but display RMSE and coverage probability of 95 per cent CIs, respectively. Bias and RMSE levels were generally consistent between Tables I, II and III, IV, respectively, or for equal and unequal variance cases. The ranges of bias were -0.0175 to 0.0143 , -0.0157 to 0.0103 and -0.0048 to 0.0089 for $n_X = n_Y = 50, 100, 1000$, respectively, with mean bias across scenarios being approximately 0.25 per cent of the AUC being estimated. Overall RMSE levels ranged from 0.0009 to 0.0346, 0.0005 to 0.0156 and 0.00003 to 0.0012 for $n_X = n_Y = 50, 100, 1000$, respectively, with corresponding mean RMSEs 0.0055, 0.0028 and 0.0003. As expected, bias and RMSE decrease with increases in sample sizes. Other relations between variations in distributional scenarios and bias, RMSE are better seen via Figures 1 and 2, respectively. Figures 1 and 2 display bias and RMSE results as a function of missingness due to LOD (0, 10, 25, 50, 75 per cent) where different lines correspond to differing levels of ME quantified by R (0.90 to 0.25 corresponding to low to high levels). Figures are for $n_X = n_Y = 100$ only, however, patterns are consistent over sample sizes and are excluded only for clarity of exposition. Figure 1 shows that the bias of $\hat{AUC}$ is no more than marginally dependent for levels of $R = 0.90$ – 0.50 , up to half of the observed variability originating from ME. Only when we consider the extreme case of $R = 0.25$, where three-fourths of overall observed variability is due to random error, do we see an increase in the magnitude of bias. Bias is also relatively unchanged as a function of per cent missing (flat lines in figures) until after correcting for 50 per cent missing. Figure 2 displays a decrease in RMSE as a function of AUC. Similar to bias, RMSE increases slightly as a function of R , until $R = 0.25$, exemplified by the different lines increasing in overall magnitude. RMSE increases slightly as a function of per cent missing (left to right) up to 50 per cent after which a sharp increase in RMSE is observed. The increase in RMSE of estimators correcting for $R = 0.90$ and 75 per cent missing substantially from the level at 50 per cent. This is driven by the bias of correcting for 75 per cent missing data at these levels and not variance. The vast majority of coverage probabilities of 95 per cent CI's for AUC in Tables V and VI are nominal or with in normal variation for a simulation with 2000 iterations. CI's seem to be conservative for low levels of discrimination with large per cent missing (75 per cent) and under covers for AUC = 0.90 with 75 per cent. Again, increasing the number of replicates per individual may help improve coverage probability. Overall, point and CI estimators perform well while correcting for various levels of LOD and ME.

5. EXAMPLE

Pre-eclampsia affects about 5 per cent of pregnancies, resulting in substantial maternal and neonatal morbidity and mortality [18]. Identifying women suffering from pre-eclampsia is a very important step in the management of the disease. Vascular endothelia growth factor (VEGf) and its soluble receptor 1 (sFLT-1) are promising biomarkers for such classification.

Apriori it was suspected that sFLT-1 and VEGf had substantial levels of ME. Subsequently, replicate samples were taken and analyzed. An additional source of error for both biomarkers were the manufacturer established LODs that each biomarker was subject to.

In a nested case control study, using two replicates per $n_X = 344$ the and $n_Y = 313$ women, we evaluated discriminatory abilities of these biomarkers. In order to apply our method we assessed the normality of biomarker levels with Q–Q plots. As Figure 3 demonstrates, we found that a log transformation resulted in a distribution that reasonably reflects our assumption of normally distributed biomarker levels. After the log transformation, we maximized the likelihood, equation (1), with respect to the normal parameters and subsequently generated the mean and covariance matrix estimators for sFLT-1 levels in of our transformed cases and controls

$$\begin{aligned} \widehat{\mu}_x &= 7.137, \quad \widehat{\mu}_y = 6.844 \\ \widehat{\Sigma}_x &= \begin{bmatrix} 0.652 & 0.628 \\ 0.628 & 0.638 \end{bmatrix} \quad \text{and} \quad \widehat{\Sigma}_y = \begin{bmatrix} 0.431 & 0.404 \\ 0.404 & 0.410 \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} \widehat{\mu}_x &= 2.065, \quad \widehat{\mu}_y = 2.16 \\ \widehat{\Sigma}_x &= \begin{bmatrix} 0.428 & 0.338 \\ 0.338 & 0.354 \end{bmatrix} \quad \text{and} \quad \widehat{\Sigma}_y = \begin{bmatrix} 0.295 & 0.219 \\ 0.219 & 0.233 \end{bmatrix} \end{aligned}$$

for VEGf levels, respectively. The covariance matrices lead to Reliability coefficient ≈ 0.97 for sFLT-1 and Reliability coefficient ≈ 0.77 for VEGf which corresponds to error variance making up roughly 3 and 23 per cent of the total biomarker variability, respectively. Fortunately in this instance, all sample measurements of sFLT-1 and VEGf registered above the LODs set by the test kit manufacturer resulting in no censored observations.

The individual MLEs above lead to an $\hat{AUC} = 0.613$ and 95 per cent CI (0.583–0.643) for the true AUC for sFLT-1 and $\hat{AUC} = 0.554$ with 95 per cent CI (0.522–0.585) for VEGf as a discriminator of women with and without pre-eclampsia. Neither biomarkers have outstanding capabilities but both are significantly better than chance when corrected for ME.

While no observations were censored, Figure 4 displays point (solid) and confident interval (dashed) estimators for AUC as if biomarker measurements had been affected at various levels of censoring do to the LODs. The bold lines in Figure 4 are based on the methods developed in Sections 2 and 3 utilizing both samples per woman. Point estimators for both sFLT-1 and VEGf are nearly unchanged correcting for increasing levels of missingness up to approximately 50 per cent missing as demonstrated by the nearly horizontal bold line. From 50 all the way to 80 per cent missing, $\hat{AUC}$ becomes only slightly more variable. Bold dashed lines gradually widen as per cent missing increases to allow for additional variability due to estimation being based on fewer observed measurements but at a rate such that VEGf remains a significant discriminator up to about 35 per cent missing and sFLT-1 is significant for up to 75 per cent missing.

Figure 4 also displays similar lines (non-bold) based on a single sample per woman and univariate LOD methods [9]. These lines illustrate that ignoring ME results in wider CIs and reduced point estimates $\hat{AUC}$ across increasing levels of the LOD in both sFLT-1 and VEGf. The degree to which estimators are affected by ignoring ME is obviously a function of the level of error being ignored. Using a single measurement of sFLT-1 results in minimal if any detriment in estimation due to the low level of error being corrected for, 3 per cent of sFLT-1 variability. However, ignoring ME in VEGf measurements would result in substantially reduced $\hat{AUC}$ and wider CI across all levels of per cent missing. These single measurement CIs for VEGf never display a discriminatory ability significantly better than chance.

For sFLT-1 investigators would have achieved adequate results using only one sample per woman even though using two replicates resulted in a more consistent estimator and narrower CI across levels of per cent missing. The low level ME present for this to be possible was unknown *a priori*. VEGf levels illustrate the other possibility, where the level of error is a significant contributor in the overall variability of biomarker levels. Here a single sample yields substantially wider CIs and a point estimate displaying less discriminatory ability than estimators based on replicates and corrects for ME.

6. DISCUSSION

Previous methods have corrected for ME and LOD individually. This is the first development that corrects the binormal ROC curve and AUC for both simultaneously. Through maximum likelihood methods we demonstrated both asymptotically unbiased point estimates and CIs with nominal coverage. This work could be extended to other summary measures of the ROC curve such that the Youden index and associated optimal cut-point and the partial area under the curve. Additionally, a natural extension to the $p = 2$ case we have demonstrated would be to perform the steps in the Appendix similarly to achieve estimates utilizing more than two replicates. However, the greatest benefit in terms of ME correction is realized from going from one observation, no ability to correct for ME, to two. Adding an additional replicate per person would improve estimates of error variability but is unlikely to be sufficient to justify increasing costs for data acquisition by 50 per cent over using two replicates.

The methods we have developed here first estimate the normal distribution parameters for diseased and non-diseased populations accounting for both ME and LOD through maximizing the appropriate likelihood function for these data. We then use standard binormal techniques and these corrected normal distribution parameters to achieve estimates for the ROC and AUC corrected for ME and LOD. This technique has the benefit of using the familiar binormal ROC curve and AUC with all the properties attached such as being invariant to monotonic transformations, but also the restrictive requirement that the data are normally distributed. As with all parametric methods, these estimators should only be used when distributional assumptions are adequately assessed and reasonably satisfied.

We assessed robustness by using gamma distributed data via simulation similar to that conducted in Section 4. The gamma distribution was chosen for its non-symmetric shape. The distribution of controls was fixed with $\alpha_Y = 4$ and $\beta_Y = 1$, the distribution of cases was adjusted to achieve AUC = 0.6, 0.7, 0.8, 0.9. The results were encouraging with the bias as a percentage of AUC averaged less than 1 per cent across sample sizes $n = 50, 100, 1000$ displaying that reasonable departures from normality do not lead to overly biased estimation. Coverage probably held up as well, averaging 97 and 96 per cent for sample sizes $n = 50, 100$, respectively, but deteriorated greatly at $n = 1000$ where coverage averaged 75 per cent, mainly do to AUC = 0.9, 51 per cent coverage. These results are to be expected as when $n = 1000$ one should be able to adequately address distributional assumptions as reasonably satisfied so this decreased coverage is unlikely to occur in practice. It has also been shown elsewhere that similar estimators are robust to small departures from normality but are unreliable for gross departures from such assumptions both [3,9,19].

Correcting for ME and LOD may help identify potentially useful discriminators of disease status that might have been overlooked if only analyzed naively and without regard to the negative impact of ME and LOD. Biomarker showing potential despite measurement difficulties may warrant future investment of resources to improve measurement procedures and technologies to realize this ability.

ACKNOWLEDGEMENTS

This research was supported by the Intramural Research Program of the Eunice Kennedy Shriver National Institutes of Health, Epidemiology Branch, Division of Epidemiology, Statistics and Prevention, National Institute for Child Health and Human Development.

Contract/grant sponsor: Intramural Research Program

Appendix

APPENDIX

The asymptotic variance of $\hat{AUC}$, σ_A^2 , can be derived via Fisher's information matrix $\mathbf{I}$ for MLE's $\hat{\mu}_X$, $\hat{\mu}_Y$, $\hat{\sigma}_X^2$, $\hat{\sigma}_Y^2$, $\hat{\sigma}_\varepsilon^2$ and the delta method. Here we can rely on similar developments for the general case using two different biomarkers. Differentiating with respect to μ_Y and σ_Y here can be written in terms of derivatives of the general log likelihood and partial derivatives of the relations between the general parameters and the parameters here

$$\frac{d \log L}{d\mu_Y} = \left(\frac{d \log L}{d\mu_1} + \frac{d \log L}{d\mu_2} \right) \frac{d\mu_1}{d\mu_Y} = \frac{d \log L}{d\mu_1} + \frac{d \log L}{d\mu_2}$$

and

$$\frac{d \log L}{d\sigma_Y} = \left(\frac{d \log L}{d\sigma_1} + \frac{d \log L}{d\sigma_2} \right) \frac{d\sigma_1}{d\sigma_Y} + \frac{d \log L}{d\rho} \frac{d\rho}{d\sigma_Y}$$

respectively, where

$$\frac{d\sigma_1}{d\sigma_Y} = \frac{1}{2}(\sigma_Y^2 + \sigma_\varepsilon^2)^{-1/2} (2\sigma_Y) = \sqrt{\rho} \quad \text{and} \quad \frac{d\rho}{d\sigma_Y} = \frac{2\sigma_Y\sigma_\varepsilon^2}{(\sigma_Y^2 + \sigma_\varepsilon^2)^2}$$

Differentiating once more yields

$$\frac{d^2 \log L}{d\mu_Y^2} = \frac{d^2 \log L}{d\mu_1^2} + \frac{d^2 \log L}{d\mu_2^2}$$

and

$$\begin{aligned} \frac{d^2 \log L}{d\sigma_Y^2} = & \left(\frac{d^2 \log L}{d\sigma_1^2} + \frac{d^2 \log L}{d\sigma_2^2} \right) \left(\frac{d\sigma_1}{d\sigma_Y} \right)^2 + \left(\frac{d \log L}{d\sigma_1} + \frac{d \log L}{d\sigma_2} \right) \\ & \times \left(\frac{d^2 \sigma_1}{d\sigma_Y^2} \right) \left(\frac{d^2 \log L}{d\rho^2} \right) \left(\frac{d\rho}{d\sigma_Y} \right)^2 + \left(\frac{d \log L}{d\rho} \right) \left(\frac{d^2 \rho}{d\sigma_Y^2} \right) \end{aligned}$$

where

$$\frac{d^2 \sigma_1}{d\sigma_Y^2} = \frac{1}{2\sqrt{\rho}} \frac{d^2 \rho}{d\sigma_Y^2} \quad \text{and} \quad \frac{d^2 \rho}{d\sigma_Y^2} = \frac{2\sigma_\varepsilon^2 (\sigma_Y^2 + \sigma_\varepsilon^2 - 4\sigma_Y)}{(\sigma_Y^2 + \sigma_\varepsilon^2)^3}$$

The first and second derivatives of the log likelihood with respect to general parameters can be substituted from equations in Perkins [15].

As our biomarker is affected by an LOD, the MLE's $\hat{\mu}_Y$ and $\hat{\sigma}_Y$ are no longer independent. Thus, we must construct a mixed term, differentiating with respect to both

$$\frac{d^2 \log L}{d\sigma_Y d\mu_Y} = \left(\frac{d^2 \log L}{d\sigma_1 d\mu_1} + \frac{d^2 \log L}{d\sigma_2 d\mu_1} + \frac{d^2 \log L}{d\sigma_1 d\mu_2} + \frac{d^2 \log L}{d\sigma_2 d\mu_2} \right) \frac{d\sigma_1}{d\sigma_Y} + \left(\frac{d^2 \log L}{d\rho d\mu_1} + \frac{d^2 \log L}{d\rho d\mu_2} \right) \frac{d\rho}{d\sigma_Y}$$

The expected values of these second derivatives also follow directly from the general bivariate case. Let us denote

$$\mathbf{I}_{ab} = \lim_{n \rightarrow \infty} \frac{-1}{n} E \left[\frac{\partial^2 \log \left(L(\vec{\theta}; \vec{z}) \right)}{\partial \theta_a \partial \theta_b} \right]$$

as an element of the Fisher information matrix. As the derivatives of the general parameter values with respect to the parameter values for the replicates (i.e. $d\sigma_1/d\sigma_Y$) do not contain any observed data, Y_j , then the expected value here leading to $\mathbf{I}_{11}$, $\mathbf{I}_{12}$ and $\mathbf{I}_{22}$ are simply the previous expected values weighted by those derivatives. For example,

$$\begin{aligned} \mathbf{I}_{22} = & E \left[\frac{d^2 \log L}{d\sigma_Y^2} \right] = \left(E \left[\frac{d^2 \log L}{d\sigma_1^2} \right] + E \left[\frac{d^2 \log L}{d\sigma_2^2} \right] \right) \left(\frac{d\sigma_1}{d\sigma_Y} \right)^2 \\ & + \left(E \left[\frac{d^2 \log L}{d\sigma_1 d\sigma_2} \right] + E \left[\frac{d^2 \log L}{d\sigma_2 d\sigma_1} \right] \right) \left(\frac{d\sigma_1}{d\sigma_Y} \right) \left(\frac{d\sigma_2}{d\sigma_Y} \right) \\ & + E \left[\frac{d^2 \log L}{d\rho^2} \right] \left(\frac{d\rho}{d\sigma_Y} \right)^2 + E \left[\frac{d^2 \log L}{d\rho d\sigma_1} \right] \left(\frac{d\rho}{d\sigma_Y} \right) \left(\frac{d\sigma_1}{d\sigma_Y} \right) \\ & + E \left[\frac{d^2 \log L}{d\rho d\sigma_2} \right] \left(\frac{d\rho}{d\sigma_Y} \right) \left(\frac{d\sigma_2}{d\sigma_Y} \right) \end{aligned}$$

REFERENCES

1. Pepe, MS. The Statistical Evaluation of Medical Tests for Classification and Prediction. Oxford University Press; New York: 2003.
2. Zhou, XH.; Obuchowski, NA.; McClish, DK. Statistical Methods in Diagnostic Medicine. Wiley; New York: 2002.
3. Schisterman EF, Faraggi D, Reiser B, Trevisan M. Statistical inference for the area under the ROC curve in the presence of random measurement error. American Journal of Epidemiology 2001;152:174–179. [PubMed: 11447052]
4. Reiser B. The effect of random measurement error on receiver operating characteristic (ROC) curves. Statistics in Medicine 2000;19:61–70. [PubMed: 10623913]
5. Hornung RW, Reed LD. Estimation of average concentration in the presence of nondetectable values. Applied Occupational and Environmental Hygiene 1990;5:46–51.
6. Haas CH, Scheff PA. Estimation of averages in truncated samples. Environmental Science and Technology 1990;24:912–919.
7. Singh A, Nocerino J. Robust estimation of mean and variance using environmental data sets with below detection limit observations. Chemometrics and Intelligent Laboratory Systems 2002;60:69–86.
8. Lynn HS. Maximum likelihood inference for left-censored HIV RNA data. Statistics in Medicine 2001;20:33–45. [PubMed: 11135346]
9. Perkins NJ, Schisterman EF, Vexler A. Receiver operating characteristic curve inference from a sample with a limit of detection. American Journal of Epidemiology 2007;165:325–333. [PubMed: 17110640]
10. Cohen AC. Estimating the mean and variance of normal populations from singly truncated and doubly truncated samples. The Annals of Mathematical Statistics 1950;21:557–569.
11. Gupta AK. Estimation of the mean and standard deviation of a normal population from a censored sample. Biometrika 1952;39:260–273.

12. Harter HL, Moore AH. Iterative maximum-likelihood estimation of the parameters of normal populations from singly and doubly censored samples. *Biometrika* 1966;53:205–213. [PubMed: 5964058]
13. Faraggi D. The effect of random measurement error on receiver operating characteristic (ROC) curves. *Statistics in Medicine* 2000;19:61–70. [PubMed: 10623913]
14. Lyles RH, Williams JK, Chuachoowong R. Correlating two viral load assays with known detection limits. *Biometrics* 2001;57:1238–1244. [PubMed: 11764265]
15. Perkins NJ. ROC inference based on best linear combination of biomarkers affected by different sources of measurement error. Dissertation. 2007
16. Perkins NJ, Schisterman EF. The Youden index and the optimal cut-point corrected for measurement error. *Biometrical Journal* 2005;47:428–441. [PubMed: 16161802]
17. Carmines EG, Zeller RA. Reliability and validity. series: quantitative applications in social sciences. Sage University Paper. 1976
18. Levine RJ, Maynard SE, Qian C, Lim KH, England LJ, Yu KF. Circulating angiogenic factors and the risk of preeclampsia. *The New England Journal of Medicine* 2004;350:672–683. [PubMed: 14764923]
19. Vexler A, Liu A, Eliseeva E, Schisterman EF. Maximum likelihood ratio tests for comparing the discriminatory ability of biomarkers subject to limit of detection. *Biometrics* 2008;64:895–903. [PubMed: 18047527]

Figure 1.

Bias of MLE's for various AUC, using two replicates on $n_X = n_Y = 100$ individuals to corrected for various levels of measurement error, as a function of the per cent of the non-diseased population missing due to an LOD under different distributional scenarios. The lines are labeled in the upper right plot by reliability index, corresponding to differing line types, or the proportion of biomarker variability to the variability of levels measured with error.

Figure 2.

Root mean-square error, RMSE, of MLE's for various AUC, using two replicates on $n_X = n_Y = 100$ individuals to corrected for various levels of measurement error, as a function of the percent of the non-diseased population missing due to an LOD under different distributional scenarios. The lines are labeled in the upper left plot by reliability index, corresponding to differing line types, or the proportion of true biomarker variability to the variability of levels measured with error.

Figure 3.
Normal Q-Q plots of log transformed sFLT-1 levels and VEGf levels in women with and without pre-eclampsia.

Figure 4.

MLEs of AUC (solid) with confidence intervals (dashed) as a function of the per cent of the controls that are missing due to an LOD for sFLT-1 and VEGf levels in women with and without pre-eclampsia. The bold lines correspond to estimators using both replicates, while the lighter lines are based on a single sample.

Table 1

Bias of the MLE for AUC, the area under the curve, based on two normally distributed replicates of biomarkers levels; equal variance case, $(\sigma_x^2, \sigma_y^2) = (1.0, 1.0)$.

AUC	Per cent mis.	R=0.90						0.75			0.50			0.25		
		$n_x = n_y = 50$														
		0	10	25	50	75	0	10	25	50	100	1000	50	100	1000	1000
0.6	0	0.0008	0.0007	0.0001	0.0001	0.0011	0.0008	0.0002	0.0002	0.0022	0.0012	0.0004	0.0086	0.0036	0.0008	0.0008
	10	0.0008	0.0005	0.0001	0.0001	0.0012	0.0006	0.0002	0.0002	0.0024	0.0010	0.0004	0.0088	0.0036	0.0009	0.0009
	25	0.0007	0.0004	0.0001	0.0001	0.0010	0.0005	0.0002	0.0002	0.0022	0.0011	0.0003	0.0103	0.0040	0.0008	0.0008
	50	0.0004	0.0004	0.0002	0.0002	0.0003	0.0005	0.0002	0.0002	0.0021	0.0007	0.0005	0.0095	0.0032	0.0011	0.0011
	75	-0.0052	-0.0020	-0.0002	-0.0002	0.0000	-0.0018	-0.0001	-0.0001	0.0035	-0.0012	0.0006	0.0088	0.0067	0.0016	0.0016
0.7	0	0.0014	0.0010	0.0000	0.0000	0.0020	0.0011	0.0001	0.0001	0.0039	0.0018	0.0003	0.0135	0.0062	0.0009	0.0009
	10	0.0015	0.0009	0.0000	0.0000	0.0020	0.0010	0.0001	0.0001	0.0038	0.0018	0.0004	0.0137	0.0063	0.0009	0.0009
	25	0.0011	0.0007	0.0001	0.0001	0.0017	0.0009	0.0002	0.0002	0.0037	0.0018	0.0003	0.0143	0.0064	0.0009	0.0009
	50	0.0004	0.0009	0.0000	0.0000	0.0009	0.0006	0.0001	0.0001	0.0034	0.0016	0.0003	0.0143	0.0070	0.0011	0.0011
	75	-0.0082	-0.0041	-0.0003	-0.0003	-0.0014	-0.0011	0.0001	0.0001	0.0044	0.0001	0.0003	0.0138	0.0103	0.0017	0.0017
0.8	0	0.0016	0.0010	0.0000	0.0000	0.0022	0.0011	0.0001	0.0001	0.0043	0.0019	0.0003	0.0125	0.0066	0.0009	0.0009
	10	0.0015	0.0010	0.0000	0.0000	0.0021	0.0011	0.0001	0.0001	0.0042	0.0020	0.0003	0.0125	0.0066	0.0008	0.0008
	25	0.0014	0.0009	0.0000	0.0000	0.0020	0.0011	0.0001	0.0001	0.0042	0.0018	0.0003	0.0128	0.0067	0.0008	0.0008
	50	0.0007	0.0007	0.0029	0.0029	0.0014	0.0007	0.0000	0.0000	0.0040	0.0020	0.0002	0.0126	0.0077	0.0010	0.0010
	75	-0.0098	-0.0062	0.0013	0.0013	-0.0025	-0.0005	0.0001	0.0001	0.0041	0.0012	0.0005	0.0107	0.0089	0.0015	0.0015
0.9	0	0.0010	0.0006	-0.0001	-0.0001	0.0014	0.0007	0.0000	0.0000	0.0026	0.0011	0.0001	0.0057	0.0036	0.0005	0.0005
	10	0.0007	0.0006	0.0006	0.0006	0.0012	0.0007	0.0000	0.0000	0.0024	0.0012	0.0001	0.0056	0.0036	0.0004	0.0004
	25	0.0006	0.0006	0.0020	0.0020	0.0010	0.0006	0.0000	0.0000	0.0023	0.0011	0.0001	0.0055	0.0035	0.0005	0.0005
	50	0.0005	0.0005	0.0031	0.0031	0.0009	0.0005	0.0089	0.0089	0.0023	0.0011	0.0001	0.0049	0.0041	0.0006	0.0006
	75	-0.0114	-0.0074	0.0022	0.0022	-0.0025	-0.0005	0.0062	0.0062	0.0023	0.0010	0.0002	0.0023	0.0036	0.0006	0.0006

$R = \sigma_x^2 + \sigma_y^2 / (\sigma_x^2 + \sigma_y^2 + 2\sigma_\epsilon^2)$, where (σ_x^2, σ_y^2) are the variances of diseased and non-diseased populations and σ_ϵ^2 is error variance.

Table II

Bias of the MLE for AUC, the area under the curve, based on two normally distributed replicates of biomarkers levels; un-equal variance case, $(\sigma_x^2, \sigma_y^2) = (0.5, 1.0)$.

AUC	Per cent mis.	R=0.90						0.75			0.50			0.25		
		$n_x = n_y = 50$														
		1000	100	10	50	100	1000	1000	100	50	1000	100	50	1000	100	50
0.6	0	0.0006	0.0010	0.0001	0.0009	0.0010	0.0002	0.0021	0.0014	0.0004	0.0004	0.0014	0.0074	0.0037	0.0036	0.0009
	10	0.0010	0.0010	0.0001	0.0012	0.0010	0.0002	0.0023	0.0014	0.0004	0.0004	0.0014	0.0076	0.0036	0.0036	0.0009
	25	0.0008	0.0009	0.0002	0.0009	0.0009	0.0002	0.0021	0.0015	0.0003	0.0003	0.0015	0.0084	0.0041	0.0041	0.0009
	50	-0.0004	0.0006	0.0002	-0.0003	0.0005	0.0002	0.0016	0.0007	0.0004	0.0004	0.0007	0.0093	0.0029	0.0029	0.0012
	75	-0.0113	-0.0120	-0.0007	-0.0040	-0.0030	-0.0001	0.0012	-0.0029	0.0005	0.0005	-0.0029	0.0047	0.0025	0.0025	0.0013
0.7	0	0.0013	0.0012	0.0001	0.0018	0.0014	0.0001	0.0037	0.0021	0.0004	0.0004	0.0021	0.0117	0.0061	0.0061	0.0009
	10	0.0013	0.0013	0.0001	0.0019	0.0014	0.0002	0.0038	0.0021	0.0004	0.0004	0.0021	0.0120	0.0061	0.0061	0.0010
	25	0.0014	0.0013	0.0001	0.0017	0.0013	0.0002	0.0036	0.0021	0.0003	0.0003	0.0021	0.0127	0.0064	0.0064	0.0009
	50	0.0003	0.0010	0.0001	0.0010	0.0009	0.0001	0.0031	0.0019	0.0004	0.0004	0.0019	0.0140	0.0063	0.0063	0.0012
	75	-0.0142	-0.0093	-0.0031	-0.0039	-0.0027	0.0001	0.0030	-0.0008	0.0002	0.0002	0.0008	0.0113	0.0087	0.0087	0.0014
0.8	0	0.0014	0.0012	0.0000	0.0020	0.0013	0.0001	0.0041	0.0021	0.0003	0.0003	0.0021	0.0110	0.0063	0.0063	0.0007
	10	0.0013	0.0012	0.0000	0.0019	0.0014	0.0001	0.0039	0.0022	0.0002	0.0002	0.0022	0.0110	0.0064	0.0064	0.0008
	25	0.0013	0.0013	0.0000	0.0018	0.0014	0.0000	0.0039	0.0022	0.0002	0.0002	0.0022	0.0116	0.0065	0.0065	0.0008
	50	0.0010	0.0011	0.0000	0.0015	0.0012	0.0000	0.0040	0.0023	0.0002	0.0002	0.0023	0.0122	0.0076	0.0076	0.0009
	75	-0.0158	-0.0125	-0.0041	-0.0040	-0.0017	0.0000	0.0039	0.0011	0.0003	0.0003	0.0011	0.0107	0.0086	0.0086	0.0012
0.9	0	0.0008	0.0007	-0.0001	0.0012	0.0008	0.0000	0.0024	0.0012	0.0001	0.0001	0.0012	0.0046	0.0033	0.0033	0.0004
	10	0.0004	0.0007	-0.0001	0.0009	0.0008	0.0000	0.0021	0.0013	0.0001	0.0001	0.0013	0.0044	0.0033	0.0033	0.0005
	25	0.0004	0.0007	-0.0001	0.0008	0.0008	0.0000	0.0021	0.0013	0.0001	0.0001	0.0013	0.0045	0.0033	0.0033	0.0005
	50	0.0006	0.0008	-0.0001	0.0009	0.0008	0.0000	0.0022	0.0013	0.0001	0.0001	0.0013	0.0043	0.0039	0.0039	0.0006
	75	-0.0175	-0.0157	-0.0048	-0.0038	-0.0011	0.0000	0.0022	0.0016	0.0001	0.0001	0.0016	0.0025	0.0039	0.0039	0.0007

$R = \sigma_x^2 + \sigma_y^2 / (\sigma_x^2 + \sigma_y^2 + 2\sigma_\epsilon^2)$, where (σ_x^2, σ_y^2) are the variances of diseased and non-diseased populations and σ_ϵ^2 is error variance.

Table III

Root mean-square error of the MLE for AUC, the area under the curve, based on two normally distributed replicates of biomarkers levels; equal variance case, $(\sigma_x^2, \sigma_y^2) = (1.0, 1.0)$.

AUC	Per cent mis.	R=0.90					0.75					0.50					0.25				
		$n_x = n_y = 50$																			
		0	10	25	50	75	0	10	25	50	75	0	10	25	50	75	0	10	25	50	75
0.6	0	0.0031	0.0016	0.0002	0.0002	0.0034	0.0018	0.0002	0.0002	0.0046	0.0023	0.0002	0.0002	0.0100	0.0043	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	10	0.0031	0.0016	0.0002	0.0002	0.0035	0.0018	0.0002	0.0002	0.0047	0.0023	0.0002	0.0002	0.0101	0.0044	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	25	0.0032	0.0017	0.0002	0.0002	0.0036	0.0018	0.0002	0.0002	0.0049	0.0024	0.0002	0.0002	0.0107	0.0045	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	50	0.0038	0.0020	0.0002	0.0002	0.0042	0.0022	0.0002	0.0002	0.0059	0.0030	0.0003	0.0003	0.0146	0.0060	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
	75	0.0122	0.0074	0.0005	0.0005	0.0093	0.0045	0.0004	0.0004	0.0120	0.0057	0.0006	0.0006	0.0327	0.0155	0.0012	0.0012	0.0012	0.0012	0.0012	0.0012
0.7	0	0.0026	0.0014	0.0001	0.0001	0.0030	0.0015	0.0002	0.0002	0.0041	0.0020	0.0002	0.0002	0.0090	0.0040	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	10	0.0026	0.0014	0.0001	0.0001	0.0030	0.0015	0.0002	0.0002	0.0041	0.0020	0.0002	0.0002	0.0092	0.0040	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	25	0.0027	0.0014	0.0001	0.0001	0.0030	0.0015	0.0002	0.0002	0.0042	0.0021	0.0002	0.0002	0.0096	0.0041	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	50	0.0030	0.0016	0.0002	0.0002	0.0034	0.0017	0.0002	0.0002	0.0048	0.0024	0.0002	0.0002	0.0116	0.0052	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
	75	0.0103	0.0064	0.0004	0.0004	0.0067	0.0031	0.0003	0.0003	0.0086	0.0042	0.0004	0.0004	0.0218	0.0107	0.0009	0.0009	0.0009	0.0009	0.0009	0.0009
0.8	0	0.0019	0.0010	0.0001	0.0001	0.0021	0.0011	0.0001	0.0001	0.0030	0.0015	0.0002	0.0002	0.0067	0.0032	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
	10	0.0019	0.0010	0.0001	0.0001	0.0021	0.0011	0.0001	0.0001	0.0030	0.0015	0.0002	0.0002	0.0068	0.0032	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
	25	0.0019	0.0010	0.0001	0.0001	0.0022	0.0011	0.0001	0.0001	0.0031	0.0015	0.0002	0.0002	0.0071	0.0033	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
	50	0.0020	0.0010	0.0001	0.0001	0.0023	0.0011	0.0001	0.0001	0.0035	0.0016	0.0002	0.0002	0.0081	0.0039	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	75	0.0092	0.0059	0.0001	0.0001	0.0046	0.0019	0.0002	0.0002	0.0052	0.0025	0.0003	0.0003	0.0128	0.0068	0.0007	0.0007	0.0007	0.0007	0.0007	0.0007
0.9	0	0.0009	0.0005	0.0000	0.0000	0.0010	0.0005	0.0001	0.0001	0.0015	0.0007	0.0001	0.0001	0.0033	0.0017	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
	10	0.0009	0.0005	0.0000	0.0000	0.0010	0.0005	0.0001	0.0001	0.0015	0.0007	0.0001	0.0001	0.0033	0.0017	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
	25	0.0009	0.0005	0.0001	0.0001	0.0011	0.0005	0.0001	0.0001	0.0016	0.0008	0.0001	0.0001	0.0034	0.0018	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
	50	0.0009	0.0005	0.0001	0.0001	0.0011	0.0005	0.0001	0.0001	0.0016	0.0008	0.0001	0.0001	0.0038	0.0020	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
	75	0.0086	0.0057	0.0001	0.0001	0.0029	0.0010	0.0001	0.0001	0.0022	0.0010	0.0001	0.0001	0.0053	0.0029	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003

$R = \sigma_x^2 + \sigma_y^2 / (\sigma_x^2 + \sigma_y^2 + 2\sigma_\epsilon^2)$, where (σ_x^2, σ_y^2) are the variances of diseased and non-diseased populations and σ_ϵ^2 is error variance.

Table IV

Root mean-square error of the MLE for AUC, the area under the curve, based on two normally distributed replicates of biomarkers levels; un-equal variance case, $(\sigma_x^2, \sigma_y^2) = (0.5, 1.0)$.

AUC	Per cent mis.	R=0.90					0.75					0.50					0.25				
		$n_x = n_y = 50$																			
		0	10	25	50	75	0	10	25	50	75	0	10	25	50	75	0	10	25	50	75
0.6	0	0.0031	0.0016	0.0002	0.0002	0.0035	0.0018	0.0002	0.0047	0.0023	0.0002	0.0047	0.0023	0.0002	0.0094	0.0043	0.0004	0.0043	0.0004	0.0043	0.0004
	10	0.0031	0.0016	0.0002	0.0002	0.0035	0.0018	0.0002	0.0047	0.0023	0.0002	0.0047	0.0023	0.0002	0.0097	0.0043	0.0004	0.0043	0.0004	0.0043	0.0004
	25	0.0032	0.0017	0.0002	0.0002	0.0036	0.0018	0.0002	0.0048	0.0024	0.0002	0.0048	0.0024	0.0002	0.0103	0.0046	0.0004	0.0046	0.0004	0.0046	0.0004
	50	0.0037	0.0020	0.0002	0.0002	0.0041	0.0022	0.0002	0.0058	0.0029	0.0003	0.0058	0.0029	0.0003	0.0144	0.0060	0.0005	0.0060	0.0005	0.0060	0.0005
	75	0.0219	0.0200	0.0025	0.0025	0.0113	0.0053	0.0004	0.0134	0.0059	0.0006	0.0134	0.0059	0.0006	0.0346	0.0156	0.0011	0.0156	0.0011	0.0156	0.0011
0.7	0	0.0027	0.0014	0.0001	0.0001	0.0030	0.0015	0.0001	0.0041	0.0020	0.0001	0.0041	0.0020	0.0001	0.0087	0.0040	0.0004	0.0040	0.0004	0.0040	0.0004
	10	0.0027	0.0014	0.0001	0.0001	0.0030	0.0015	0.0001	0.0041	0.0020	0.0001	0.0041	0.0020	0.0001	0.0088	0.0040	0.0004	0.0040	0.0004	0.0040	0.0004
	25	0.0027	0.0014	0.0001	0.0001	0.0030	0.0016	0.0001	0.0042	0.0021	0.0001	0.0042	0.0021	0.0001	0.0093	0.0041	0.0004	0.0041	0.0004	0.0041	0.0004
	50	0.0029	0.0015	0.0001	0.0001	0.0033	0.0017	0.0002	0.0047	0.0023	0.0002	0.0047	0.0023	0.0002	0.0116	0.0051	0.0004	0.0051	0.0004	0.0051	0.0004
	75	0.0137	0.0102	0.0022	0.0022	0.0073	0.0036	0.0002	0.0090	0.0040	0.0004	0.0090	0.0040	0.0004	0.0227	0.0109	0.0008	0.0109	0.0008	0.0109	0.0008
0.8	0	0.0019	0.0010	0.0001	0.0001	0.0022	0.0011	0.0001	0.0031	0.0015	0.0001	0.0031	0.0015	0.0001	0.0066	0.0032	0.0003	0.0032	0.0003	0.0032	0.0003
	10	0.0019	0.0010	0.0001	0.0001	0.0022	0.0011	0.0001	0.0031	0.0015	0.0001	0.0031	0.0015	0.0001	0.0066	0.0032	0.0003	0.0032	0.0003	0.0032	0.0003
	25	0.0019	0.0010	0.0001	0.0001	0.0022	0.0011	0.0001	0.0031	0.0015	0.0001	0.0031	0.0015	0.0001	0.0069	0.0033	0.0003	0.0033	0.0003	0.0033	0.0003
	50	0.0020	0.0010	0.0001	0.0001	0.0023	0.0011	0.0001	0.0034	0.0016	0.0001	0.0034	0.0016	0.0001	0.0080	0.0038	0.0003	0.0038	0.0003	0.0038	0.0003
	75	0.0128	0.0106	0.0028	0.0028	0.0054	0.0025	0.0001	0.0056	0.0024	0.0002	0.0056	0.0024	0.0002	0.0130	0.0068	0.0005	0.0068	0.0005	0.0068	0.0005
0.9	0	0.0009	0.0005	0.0000	0.0000	0.0011	0.0005	0.0001	0.0015	0.0008	0.0001	0.0015	0.0008	0.0001	0.0032	0.0017	0.0002	0.0017	0.0002	0.0017	0.0002
	10	0.0009	0.0005	0.0000	0.0000	0.0011	0.0005	0.0001	0.0016	0.0008	0.0001	0.0016	0.0008	0.0001	0.0033	0.0017	0.0002	0.0017	0.0002	0.0017	0.0002
	25	0.0009	0.0005	0.0000	0.0000	0.0011	0.0005	0.0001	0.0016	0.0008	0.0001	0.0016	0.0008	0.0001	0.0034	0.0017	0.0002	0.0017	0.0002	0.0017	0.0002
	50	0.0010	0.0005	0.0000	0.0000	0.0011	0.0006	0.0001	0.0017	0.0008	0.0001	0.0017	0.0008	0.0001	0.0038	0.0020	0.0002	0.0020	0.0002	0.0020	0.0002
	75	0.0127	0.0117	0.0033	0.0033	0.0039	0.0017	0.0001	0.0027	0.0010	0.0001	0.0027	0.0010	0.0001	0.0052	0.0029	0.0003	0.0029	0.0003	0.0029	0.0003

$R = \sigma_x^2 + \sigma_y^2 / (\sigma_x^2 + \sigma_y^2 + 2\sigma_\epsilon^2)$, where (σ_x^2, σ_y^2) are the variances of diseased and non-diseased populations and σ_ϵ^2 is error variance.

Table V

AUC	Per cent mis.	R=0.90				0.75			0.50			0.25		
		$n_x = n_y = 50$				50	100	1000	50	100	1000	50	100	1000
		100	1000											
0.6	0	0.9590	0.9420	0.9495	0.9575	0.9445	0.9470	0.9605	0.9515	0.9435	0.9735	0.9595	0.9420	
	10	0.9605	0.9470	0.9480	0.9620	0.9510	0.9500	0.9650	0.9570	0.9550	0.9760	0.9640	0.9525	
	25	0.9740	0.9715	0.9700	0.9695	0.9630	0.9645	0.9665	0.9650	0.9595	0.9770	0.9595	0.9550	
	50	0.9860	0.9875	0.9835	0.9795	0.9755	0.9770	0.9675	0.9620	0.9655	0.9745	0.9570	0.9490	
	75	0.9520	0.9585	0.9910	0.9710	0.9810	0.9885	0.9375	0.9685	0.9595	0.7555	0.7885	0.8920	
0.7	0	0.9560	0.9470	0.9490	0.9550	0.9515	0.9480	0.9570	0.9525	0.9440	0.9715	0.9565	0.9420	
	10	0.9555	0.9485	0.9505	0.9565	0.9520	0.9475	0.9605	0.9545	0.9480	0.9750	0.9635	0.9505	
	25	0.9615	0.9580	0.9605	0.9615	0.9605	0.9575	0.9640	0.9625	0.9585	0.9715	0.9660	0.9490	
	50	0.9800	0.9810	0.9800	0.9750	0.9725	0.9715	0.9665	0.9625	0.9605	0.9735	0.9625	0.9430	
	75	0.9510	0.9585	0.9935	0.9665	0.9800	0.9815	0.9430	0.9550	0.9535	0.8990	0.9250	0.9075	
0.8	0	0.9560	0.9480	0.9510	0.9510	0.9505	0.9515	0.9560	0.9470	0.9450	0.9710	0.9580	0.9380	
	10	0.9495	0.9440	0.9505	0.9490	0.9485	0.9495	0.9575	0.9495	0.9425	0.9750	0.9605	0.9400	
	25	0.9545	0.9475	0.9475	0.9500	0.9510	0.9495	0.9560	0.9555	0.9510	0.9735	0.9590	0.9395	
	50	0.9700	0.9685	0.9615	0.9635	0.9645	0.9640	0.9645	0.9645	0.9540	0.9705	0.9680	0.9405	
	75	0.9485	0.9605	0.9880	0.9650	0.9720	0.9760	0.9290	0.9455	0.9485	0.8940	0.9065	0.8840	
0.9	0	0.9520	0.9530	0.9535	0.9520	0.9505	0.9515	0.9585	0.9470	0.9460	0.9655	0.9590	0.9305	
	10	0.9370	0.9390	0.9410	0.9435	0.9415	0.9415	0.9560	0.9435	0.9415	0.9625	0.9590	0.9295	
	25	0.9365	0.9370	0.9185	0.9400	0.9390	0.9405	0.9550	0.9485	0.9410	0.9650	0.9555	0.9295	
	50	0.9460	0.9455	0.9030	0.9460	0.9410	0.7600	0.9670	0.9605	0.9465	0.9565	0.9595	0.9350	
	75	0.9370	0.9515	0.9685	0.9525	0.9630	0.9000	0.9100	0.9230	0.9350	0.8700	0.8700	0.3370	

$R = \sigma_x^2 + \sigma_y^2 / (\sigma_x^2 + \sigma_y^2 + 2\sigma_\epsilon^2)$, where (σ_x^2, σ_y^2) are the variances of diseased and non-diseased populations and σ_ϵ^2 is error variance.

Table VI

Coverage probability of 95 per cent confidence intervals for AUC, the area under the curve, based on two normally distributed replicates of biomarkers levels; un-equal variance case, $(\sigma_x^2, \sigma_y^2) = (0.5, 1.0)$.

AUC	Per cent mis.	R=0.90				0.75				0.50				0.25			
		$n_x = n_y = 50$	100	1000		50	100	1000		50	100	1000		50	100	1000	
0.6	0	0.9575	0.9420	0.9500		0.9545	0.9455	0.9470		0.9595	0.9490	0.9400		0.9725	0.9600	0.9430	
	10	0.9565	0.9445	0.9495		0.9535	0.9460	0.9465		0.9605	0.9535	0.9490		0.9725	0.9655	0.9530	
	25	0.9685	0.9615	0.9625		0.9660	0.9595	0.9595		0.9640	0.9570	0.9555		0.9735	0.9605	0.9530	
	50	0.9890	0.9865	0.9835		0.9765	0.9745	0.9770		0.9610	0.9560	0.9590		0.9725	0.9570	0.9455	
	75	0.8690	0.8365	0.9735		0.9565	0.9805	0.9905		0.9070	0.9545	0.9950		0.6795	0.6705	0.7995	
0.7	0	0.9565	0.9445	0.9490		0.9510	0.9445	0.9475		0.9570	0.9495	0.9425		0.9695	0.9580	0.9385	
	10	0.9495	0.9405	0.9440		0.9485	0.9405	0.9425		0.9555	0.9505	0.9435		0.9710	0.9630	0.9450	
	25	0.9560	0.9500	0.9520		0.9560	0.9490	0.9465		0.9565	0.9570	0.9485		0.9710	0.9610	0.9405	
	50	0.9815	0.9770	0.9740		0.9665	0.9680	0.9665		0.9650	0.9560	0.9495		0.9715	0.9625	0.9380	
	75	0.9360	0.9305	0.9800		0.9635	0.9790	0.9835		0.9405	0.9575	0.9680		0.8560	0.8790	0.8990	
0.8	0	0.9500	0.9455	0.9550		0.9490	0.9480	0.9505		0.9530	0.9450	0.9450		0.9680	0.9605	0.9350	
	10	0.9410	0.9355	0.9335		0.9380	0.9375	0.9365		0.9485	0.9455	0.9385		0.9690	0.9580	0.9320	
	25	0.9395	0.9355	0.9345		0.9390	0.9370	0.9350		0.9515	0.9485	0.9375		0.9740	0.9595	0.9325	
	50	0.9580	0.9585	0.9530		0.9545	0.9540	0.9470		0.9600	0.9535	0.9415		0.9650	0.9650	0.9355	
	75	0.9355	0.9430	0.9775		0.9590	0.9670	0.9675		0.9140	0.9335	0.9410		0.8830	0.8735	0.5090	
0.9	0	0.9490	0.9460	0.9510		0.9495	0.9470	0.9495		0.9560	0.9515	0.9460		0.9630	0.9585	0.9280	
	10	0.9305	0.9230	0.9250		0.9305	0.9260	0.9280		0.9455	0.9355	0.9330		0.9610	0.9545	0.9280	
	25	0.9205	0.9165	0.9070		0.9250	0.9210	0.9215		0.9470	0.9355	0.9315		0.9610	0.9555	0.9265	
	50	0.9300	0.9255	0.9110		0.9310	0.9250	0.9210		0.9645	0.9445	0.9325		0.9555	0.9565	0.9280	
	75	0.9110	0.9155	0.9320		0.9400	0.9510	0.9475		0.8790	0.9020	0.9565		0.8390	0.8220	0.7735	

$R = \sigma_x^2 + \sigma_y^2 / (\sigma_x^2 + \sigma_y^2 + 2\sigma_\epsilon^2)$, where (σ_x^2, σ_y^2) are the variances of diseased and non-diseased populations and σ_ϵ^2 is error variance.