Maximum Likelihood Ratio Tests for Comparing the Discriminatory Ability of Biomarkers Subject to Limit of Detection

Albert Vexler,[1,2,*] **Aiyi Liu,**[1] **Ekaterina Eliseeva**[1,3] **and Enrique F. Schisterman**[1]

[1] Division of Epidemiology, Statistics and Prevention Research, NICHD, NIH, Rockville, MD 20852, U.S.A.

[2] Department of Biostatistics, The State University of New York, Farber Hall, Buffalo, NY 14214, U.S.A.

[3] The University of Chicago, Chicago, IL, U.S.A.

Summary. In this paper, we consider comparing the areas under correlated receiver operating characteristic (ROC) curves of diagnostic biomarkers whose measurements are subject to a limit of detection (LOD), a source of measurement error from instruments' sensitivity in epidemiological studies. We propose and examine the likelihood ratio tests with operating characteristics that are easily obtained by classical maximum likelihood methodology.

Key words: Area under curve (AUC); Censoring; Hypothesis testing; Limit of detection (LOD); Maximum likelihood; Receiver operating characteristics (ROC).

* *email:* avexler@buffalo.edu

1. Introduction

Receiver operating characteristic (ROC) curve is a well accepted statistical tool for evaluating the discriminatory ability of biomarkers (e.g., Shapiro, 1999). An ROC curve plots the true positive rates of a biomarker versus its false positive rates for various thresholds of the test result. It is a convenient way to compare diagnostic biomarkers since the ROC curve places tests on the same scale where they can be compared for accuracy.

The area under the ROC curve (AUC) is a common index of the diagnostic performance of a biomarker. Bamber (1975) showed that AUC = Pr(X > Y), where X and Y represent values of the biomarker from diseased and healthy populations, respectively. Obviously, the closer the AUC is to one, the better the diagnostic accuracy of the biomarker. In a parametric setting the AUCs can generally be expressed as a function of unknown parameters and thus can be evaluated via estimation of these parameters. Nonparametric estimation of the AUC has also been well addressed in the biostatistical and epidemiological literature. However, the test-scores of biomarkers are frequently associated with measurement error, and in this paper we focus on measurement errors due to the limit of detection (LOD).

The LOD is a source of bias in many experiments and is usually caused by the limitation of instruments in measuring very high or low concentrations (e.g., Lyles *et al*, 2001; Lubin *et al*, 2004; Schisterman *et al*, 2006; Vexler *et al*, 2006; Mumford *et al*, 2006). This inability to accurately determine values of biomarkers introduces bias in the analysis of data from such experiments. For example, biomarkers for polychlorinated biphenyl (PCB), which are associated with endometriosis (Louis *et al*, 2005), are limited by instrument sensitivity (e.g., Lubin *et al*, 2004). The LOD issue can be considered as a problem of censored data analysis (e.g., Vexler *et al*, 2006). Perkins *et al.* (2006) as well as Mumford *et al.* (2006) have proposed methods for estimation of ROC curves based on samples with LOD.

Often it is necessary to determine whether a biomarker has satisfactory accuracy in correctly discriminating between cases and controls, e.g. testing for $AUC = 0.5$ (i.e. a biomarker has no discriminatory ability), or whether one biomarker has better diagnostic accuracy than another (e.g., Molodianovitch *et al*, 2006). This can be achieved by comparing the AUCs of these biomarkers. The present paper addresses these issues when the measurements of the biomarkers are subject to a limit of detection. We investigate the maximum likelihood ratio test (MLRT), utilizing the likelihood function proposed by Lyles *et al.* (2001). Operating characteristics of the proposed tests (e.g. significance level and power) can be obtained from classical results of the maximum likelihood method.

The paper is organized as follows. Section 2 introduces the MLRT for comparing AUCs. Section 3 presents Monte Carlo simulation results. In section 4, we apply the proposed tests to data from two studies to evaluate the AUCs of several biomarkers, with some concluding remarks in section 5. One example is from a study conducted in Birmingham, Alabama to investigate whether intrauterine inflammation is associated with neuron developmental abnormalities in early childhood, so that certain educational methods for improvement will be utilized. In this example the levels of intrauterine inflammation biomarkers are observed only if they are above the detection limits. Another example uses data from a study of atherosclerotic coronary heart disease to test for discriminatory ability of several biomarkers. This study sampled residents of Niagara and Erie counties in New York who were between the ages of 35 and 79. Adults between the ages of 35 and 65 were randomly selected using the New York State Department of Motor Vehicles drivers' licenses rolls. Individuals between 65 and 79 years of age were sampled randomly from the Health Care Financing Administration database. A cohort of 942 individuals consisted of 143 people with myocardial infarction (cases) and 799 controls. The purpose of the study was to determine whether biomarkers that measure individuals' oxidative stress and antioxidant status are good at determining an individual's disease status.

2. Maximum Likelihood Ratio Tests

2.1 *Test Based on Complete Data*

Let X_k and Y_k represent the values of biomarker $k(=1,2)$ associated with a diseased (X) and healthy (Y) population, respectively, and $\{x_{k1},\ldots,x_{kn}\}$ and $\{y_{k1},\ldots,y_{km}\}$ be the corresponding test-scores. Suppose the independent vectors $(x_{1i},x_{2i})^T$ follow a normal distribution

$$(x_{1i},x_{2i})^T \sim N\left((\mu_{x_1},\mu_{x_2})^T, \begin{bmatrix} \sigma_{x_1}^2 & \rho_x\sigma_{x_1}\sigma_{x_2} \\ \rho_x\sigma_{x_1}\sigma_{x_2} & \sigma_{x_2}^2 \end{bmatrix}\right),\ i=1,\ldots,n,$$

and similarly,

$$(y_{1j},y_{2j})^T \sim N\left((\mu_{y_1},\mu_{y_2})^T, \begin{bmatrix} \sigma_{y_1}^2 & \rho_y\sigma_{y_1}\sigma_{y_2} \\ \rho_y\sigma_{y_1}\sigma_{y_2} & \sigma_{y_2}^2 \end{bmatrix}\right),\ j=1,\ldots,m.$$

Following Bamber (1975), the AUCs of the biomarkers are $AUC_1 = P(X_1 > Y_1)$ and $AUC_2 = P(X_2 > Y_2)$, respectively.

In this section, we formally consider testing hypothesis

$$H_0: AUC_1 = AUC_2 \text{ versus } H_1: AUC_1 \neq AUC_2. \tag{1}$$

It is clear that $AUC_k = \Phi\left\{(\mu_{x_k}-\mu_{y_k})/\sqrt{\sigma_{x_k}^2+\sigma_{y_k}^2}\right\}$, $k=1,2$, and therefore

$$AUC_1 = AUC_2 \quad \text{iff} \quad \mu_{x_1} = \frac{(\mu_{x_2}-\mu_{y_2})(\sigma_{x_1}^2+\sigma_{y_1}^2)^{1/2}}{(\sigma_{x_2}^2+\sigma_{y_2}^2)^{1/2}} + \mu_{y_1}.$$

In a simple case, where all the parameters are known and there is no measurement error, i.e. (X_1,X_2) and (Y_1,Y_2) are observed completely, we can utilize the classical maximum likelihood ratio test for testing H_0. To this end, note that under H_1 and H_0 the likelihood function has the form

$$\prod_{\substack{i=1,\ldots,n\\ j=1,\ldots,m}} f(x_{1i}, x_{2i}, y_{1j}, y_{2j}; \Theta_X^{H_1}, \Theta_Y^{H_1}), \quad \prod_{\substack{i=1,\ldots,n\\ j=1,\ldots,m}} f(x_{1i}, x_{2i}, y_{1j}, y_{2j}; \Theta_X^{H_0}, \Theta_Y^{H_0}),$$

respectively, where the vectors of parameters $\Theta_X^{H_1}, \Theta_Y^{H_1}, \Theta_X^{H_0}, \Theta_Y^{H_0}$ are

$$\Theta_X^{H_1} = (\mu_{x_1}, \mu_{x_2}, \sigma_{x_1}^2, \sigma_{x_2}^2, \rho_x), \ \Theta_Y^{H_1} = (\mu_{y_1}, \mu_{y_2}, \sigma_{y_1}^2, \sigma_{y_2}^2, \rho_y),$$
$$\Theta_X^{H_0} = \left(\frac{(\mu_{x_2} - \mu_{y_2})(\sigma_{x_1}^2 + \sigma_{y_1}^2)^{1/2}}{(\sigma_{x_2}^2 + \sigma_{y_2}^2)^{1/2}} + \mu_{y_1}, \mu_{x_2}, \sigma_{x_1}^2, \sigma_{x_2}^2, \rho_x \right), \Theta_Y^{H_0} = \Theta_Y^{H_1},$$

and the density function f is $f(x_1, x_2, y_1, y_2; \Theta_X, \Theta_Y) = \phi(x_1, x_2; \Theta_X)\phi(y_1, y_2; \Theta_Y)$, where, with $\Theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$,

$$\phi(u, v; \Theta) = \frac{1}{2\pi\theta_3\theta_4\sqrt{1-\theta_5^2}}$$
$$\times \exp\left[-\frac{1}{2}\left(\frac{1}{1-\theta_5^2}\right) \left\{ \frac{(u-\theta_1)^2}{\theta_3^2} - 2\theta_5 \frac{(u-\theta_1)}{\theta_3}\frac{(v-\theta_2)}{\theta_4} + \frac{(v-\theta_2)^2}{\theta_4^2} \right\} \right].$$

Therefore the classical likelihood ratio test-statistic is

$$z = \prod_{\substack{i=1,\ldots,n\\ j=1,\ldots,m}} \frac{f(x_{1i}, x_{2i}, y_{1j}, y_{2j}; \Theta_X^{H_1}, \Theta_Y^{H_1})}{f(x_{1i}, x_{2i}, y_{1j}, y_{2j}; \Theta_X^{H_0}, \Theta_Y^{H_0})}.$$

Thus, we reject the null hypothesis iff $z > z_\alpha$, where the threshold z_α corresponds to Type I error α. It is clear that this test is the most powerful unbiased test; see e.g. Lehmann (1997).

When the parameters $\Theta_X^{H_1}$, $\Theta_Y^{H_1}$, $\Theta_X^{H_0}$, $\Theta_Y^{H_0}$ are unknown, Molodianovitch *et al.* (2006) propose the transformed normal approach by normalizing data through transformation and then applying the parametric test proposed by Wieand *et al.* (1989) in order to test for hypothesis (1). (This test is based on confidence intervals of AUCs (e.g., Reiser and Faraggi, 1997). We will investigate this method in detail in Section 4.2.) Alternatively, we can apply the maximum likelihood estimation and obtain the test-statistic

$$z = \frac{\sup_{\bar{\mu}_{x_1},\bar{\mu}_{x_2},\bar{\mu}_{y_1},\bar{\mu}_{y_2},\bar{\sigma}^2_{x_1},\bar{\sigma}^2_{x_2},\bar{\sigma}^2_{y_1},\bar{\sigma}^2_{y_2},\bar{\rho}_x,\bar{\rho}_y} \prod_{\substack{i=1,\dots,n\\ j=1,\dots,m}} f(x_{1i}, x_{2i}, y_{1j}, y_{2j}; \overline{\Theta}_X^{H_1}, \overline{\Theta}_Y^{H_1})}{\sup_{\bar{\mu}_{x_2},\bar{\mu}_{y_1},\bar{\mu}_{y_2},\bar{\sigma}^2_{x_1},\bar{\sigma}^2_{x_2},\bar{\sigma}^2_{y_1},\bar{\sigma}^2_{y_2},\bar{\rho}_x,\bar{\rho}_y} \prod_{\substack{i=1,\dots,n\\ j=1,\dots,m}} f(x_{1i}, x_{2i}, y_{1j}, y_{2j}; \overline{\Theta}_X^{H_0}, \overline{\Theta}_Y^{H_0})},$$

where

$$\overline{\Theta}_X^{H_1} = (\bar{\mu}_{x_1}, \bar{\mu}_{x_2}, \bar{\sigma}^2_{x_1}, \bar{\sigma}^2_{x_2}, \bar{\rho}_x),\ \overline{\Theta}_Y^{H_1} = (\bar{\mu}_{y_1}, \bar{\mu}_{y_2}, \bar{\sigma}^2_{y_1}, \bar{\sigma}^2_{y_2}, \bar{\rho}_y),$$
$$\overline{\Theta}_X^{H_0} = \left(\frac{(\bar{\mu}_{x_2} - \bar{\mu}_{y_2})(\bar{\sigma}^2_{x_1} + \bar{\sigma}^2_{y_1})^{1/2}}{(\bar{\sigma}^2_{x_2} + \bar{\sigma}^2_{y_2})^{1/2}} + \bar{\mu}_{y_1}, \bar{\mu}_{x_2}, \bar{\sigma}^2_{x_1}, \bar{\sigma}^2_{x_2}, \bar{\rho}_x \right), \overline{\Theta}_Y^{H_0} = \overline{\Theta}_Y^{H_1}.$$

It is well known (e.g., Lehmann, 1997) that under H_0, the statistic $2 \log z$ asymptotically has a χ^2_1 distribution and therefore the threshold z_α can be easily obtained from $Pr(z > z_\alpha) = \alpha$, as $n, m \to \infty$. Moreover, this test is asymptotically most powerful (e.g. Choi *et al*, 1996).

2.2 *Test Based on Data Subject to Limit of Detection*

If measurements of the biomarkers are subject to a limit of detection, then instead of observing x_{1i}, x_{2i}, y_{1j}, y_{2j} we have

$$x'_{ki} = \begin{cases} x_{ki}, & \text{if } x_{ki} \ge d_x; \\ \text{NA (Not Available)}, & x_{ki} < d_x, \end{cases} \qquad y'_{kj} = \begin{cases} y_{kj}, & \text{if } y_{kj} \ge d_y; \\ \text{NA}, & y_{kj} < d_y, \end{cases}$$

where $k = 1, 2,\ i = 1, \dots, n,\ j = 1, \dots, m$ and d_x, d_y are the values of the LOD (e.g. Lynn, 2001; Lubin *et al*, 2004; Schisterman *et al*, 2006; Vexler *et al*, 2006; Mumford *et al*, 2006). In the present paper we assume, without loss of generality, that $d_x = d_y = d$ and d is known (if d is unknown, it can be easily estimated, for example, by $\min_{i,j,k}\{x_{ki}, y_{kj}\}$). We can still obtain the MLRT-statistic based on the left-censored data. Following Lyles *et al.* (2001), write the likelihood functions based on $X' = \{x'_{1i}, x'_{2i}\}_{i=1}^n$ and $Y' = \{y'_{1j}, y'_{2j}\}_{j=1}^m$ as $L\left(X'; \Theta_X^{H_1}\right)$ and $L\left(Y'; \Theta_Y^{H_1}\right)$, respectively, that are formally defined in Appendix A.

Thus the MLRT statistic is given by

$$z^{(LOD)} = \frac{\sup_{\bar{\mu}_{x_1},\bar{\mu}_{x_2},\bar{\mu}_{y_1},\bar{\mu}_{y_2},\bar{\sigma}^2_{x_1},\bar{\sigma}^2_{x_2},\bar{\sigma}^2_{y_1},\bar{\sigma}^2_{y_2},\bar{\rho}_x,\bar{\rho}_y} L\left(X';\overline{\Theta}_X^{H_1}\right) L\left(Y';\overline{\Theta}_Y^{H_1}\right)}{\sup_{\bar{\mu}_{x_2},\bar{\mu}_{y_1},\bar{\mu}_{y_2},\bar{\sigma}^2_{x_1},\bar{\sigma}^2_{x_2},\bar{\sigma}^2_{y_1},\bar{\sigma}^2_{y_2},\bar{\rho}_x,\bar{\rho}_y} L\left(X';\overline{\Theta}_X^{H_0}\right) L\left(Y';\overline{\Theta}_Y^{H_0}\right)}.$$

Subsequently, the test threshold z_α can be obtained by the MLRT's asymptotic result: $2\log z^{(LOD)} \sim \chi^2_1$ as $n, m \to \infty$ and d is fixed.

Remark 1: Numerical calculations. Note that, applying statistical software such as R, SPlus, etc. allows us to calculate test-statistics z and $z^{(LOD)}$ without using closed forms of the estimators of the unknown parameters. A schematic example of programming in R is available upon request from the first author.

Remark 2: Transformed normal approach. The proposed method is based on the MLRT technique and hence the parametric assumptions regarding the data points are required. In order to relax the normal distribution assumptions, following Molodianovitch *et al.* (2006) we can fit the data to a Box-Cox power transformation model to better achieve normality and then test for (1). Note that, Molodianovitch *et al.* (2006) have concluded that the transformed normal approach is efficient and robust when AUCs are compared. We present a modification of the proposed test in Appendix B.

3. Simulation

We conducted Monte Carlo simulations to examine the performance of the proposed method. To this end, we generated values of $\{x_{1i}, i = 1, \ldots, n\}$ from the normal distribution with mean μ_{x_1} and variance 1, and $\{x_{2i} = ax_{1i} + \varepsilon_i, i = 1, \ldots, n\}$, where the i.i.d. random variables $\varepsilon_i \sim N(0, 1)$. Similarly, $y_{1j} \sim N(1, 0.5^2)$ and $y_{2j} = by_{1j} + \varepsilon_j$ $(j = 1, \ldots, m)$ were generated.

Hence, $\rho_x = a\sigma_{x_1} / \left(1 + a^2\sigma_{x_1}^2\right)^{1/2}$, $\sigma_{x_2} = \left(a^2\sigma_{x_1}^2 + 1\right)^{1/2}$, $\sigma_{y_2} = \left(b^2\sigma_{y_1}^2 + 1\right)^{1/2}$, and $\rho_y = b\sigma_{y_1} / \left(1 + b^2\sigma_{y_1}^2\right)^{1/2}$, where a and b are specified below.

Significance Level of the Test. Setting $a = 0.7$, $b = 0.5$, $\mu_{x_1} = 1.274$, we have $AUC_1 = AUC_2 = 0.597$. For each value of $d = -3$, -1, -0.5, 0, 0.5, and 0.75 we generated 10,000 samples of $\{x_{1i}, x_{2i}, i = 1, \dots, n\}$ and $\{y_{1j}, y_{2j}, j = 1, \dots, m\}$. Based on the generated samples, in each repetition we calculated the values of the test-statistic $z^{(LOD)}$.

[Table 1 about here.]

In Table 1 we present the Monte Carlo estimation of Type I error, where the test thresholds $2\log z_\alpha$ are 3.84 and 6.63. These thresholds correspond to $\Pr(\xi > 3.84) = 0.05$ and $\Pr(\xi > 6.63) = 0.01$, where $\xi \sim \chi_1^2$. Table 1 also provides the theoretical proportion of the number of observations of X_1, X_2, Y_1, and Y_2 that are below the LOD value d. As can be seen, asymptotically, the Type I error of the proposed test can be obtained from the χ_1^2 distribution of the $2\log z^{(LOD)}$ statistic. However, if, for example $n = m = 150$ and $d = 0.75$, (in which case about 60% of Y_2's are not observed numerically), this assumption is dubious. (Note that for Type I error α we can assume for this simulation $CI = \alpha \pm 1.96\,\{\alpha(1-\alpha)/10000\}^{1/2}$.)

Power of the test. Here we examine the power of the test for situations where $\{AUC_1 = 0.5, AUC_2 = 0.6\}$ and $\{AUC_1 = 0.6, AUC_2 = 0.9\}$. For the first case, we set $\mu_{x_1} = 1.3$, $a = 0.5$, $b = -1.5$, and for the second $\mu_{x_1} = 1$, $a = 0.7$, $b = 0.3$. For both cases $n = m = 150$. Table 2 displays the Monte Carlo estimation of the test's power for different values of d. Obviously, the power of the test is dependent on the proportion of X_1's, X_2's, Y_1's, and Y_2's below d.

[Table 2 about here.]

Table 2 demonstrates the high values of the power even in the situation where AUC_1 is close to AUC_2 and the proportions of the biomarker values below d is high ($d = 0, 0.75$).

Robustness. The simulations thus far assume that the samples follow normal distributions. In order to illustrate the robustness of our method, we performed the following Monte Carlo simulations. Suppose that, instead of following normal distributions, the diagnostic markers satisfy $x_{2i} = 0.7x_{1i}^{(df)} + \varepsilon_i^{(fd)}$, $y_{2j} = 0.3y_{1j}^{(df)} + \varepsilon_j^{(fd)}$, where $x_1^{(df)}, y_1^{(df)}$ and $\varepsilon^{(df)}$ are independent identically t-distributed random variables with df degrees of freedom, mean 0 and variance 1, $1 \leq i, j \leq 150$. Thus, $AUC_1 = AUC_2 = 0.5$. Here we ran 10000 repetitions of the sample (X', Y') at each $df = 5, 10, 15$ and $d = -3, -1, 0$ (d is the value of LOD). We examined the significance level of the proposed test given the uncorrected distributional assumption. Table 3 corresponds to the case when we expect for the Type I error to be 0.05 (the test threshold $2 \log z_\alpha$ is 3.84).

[Table 3 about here.]

From these results we conclude that the proposed method is reasonable even when the distributional assumptions do not exactly satisfy normality. However, the accuracy of the expected significance level is poor when $d = 0$ (about 50% of the data are below the detection limit). In contrast, Table 1 indicates that under the corrected distributional assumption this proportion of observations below LOD is not critical.

Imputation method. Conventional approaches to dealing with data below LOD include omission, resulting in a truncated data set, and imputation with a constant, such as d or a fraction thereof (e.g., $d/2$, $d/\sqrt{2}$); or the observed values may be used directly or indirectly (e.g., Lubin *et al*, 2004; Schisterman *et al*, 2006). Perkins *et al.* (2006) showed that the imputation method can lead to biased parametric/nonparametric estimation of AUCs. Here we report results of the Monte Carlo simulation corresponding to Table 1, where the test based on confidence intervals (e.g., Reiser and Faraggi, 1997; Wieand *et al*, 1989; for details, see Section 4.2) has been calculated for observations

$$x''_{ki} = \begin{cases} x_{ki}, & x_{ki} \geq d; \\ Imp, & x_{ki} < d, \end{cases} \qquad y''_{kj} = \begin{cases} y_{kj}, & y_{kj} \geq d_y; \\ Imp, & y_{kj} < d, \end{cases}$$

$k = 1, 2,\ i = 1, \ldots, n = 150,\ j = 1, \ldots, m = 150$ and $Imp = d/2, d/\sqrt{2}$.

[Table 4 about here.]

In contrast with Table 1, Table 4 demonstrates that when $d = 0, -3, 0.75$ these conventional approaches should not be recommended. (Investigation of the test based on the samples ignoring the NA values and the nonparametric test (Wieand *et al*, 1989) based on $x''_{ki}, y''_{kj}, k = 1, 2,\ i = 1, \ldots, 150,\ j = 1, \ldots, 150$ led to similar conclusion.)

4. Examples

We exemplify the proposed method with data from the two studies briefly described in the introduction.

4.1 *The IQ Study*

Here we examine whether biomarker IL8 has the ability to discriminate between low and high levels of IQ. The data includes 369 subjects. The IQ indicator full-scale IQ (FSIQ) has values ranging from 46 to 118 with an average equal to 82.57. We split our data into two populations, where population A includes those with IQ less than 82.57 and population B includes those with IQ greater than 82.57. We associate biomarker IL8 with both populations separately. Denote X, Y as biomarker values related to population A and B, respectively. The total number of Xs is 189 and the total number of Ys is 180. According to the instrument manual, the LOD for IL8 is $d = 3.2$, yielding the numbers of NAs to be 95 and 108 for X and Y, respectively. The logarithmic values of the biomarkers are used in order to better achieve normality. The empirical histograms of the log-transformed biomarker corresponding to high and low levels of IQ are depicted in Figure 1.

[Figure 1 about here.]

Under the assumption that $\log X$ and $\log Y$ have normal distributions, applying the maximum likelihood estimation proposed by Vexler *et al.* (2006) (or estimation based on censored data, see for example Gupta (1952)) leads to estimated mean of $\log X$ and $\log Y$ as 1.02 and 0.38, respectively. The corresponding standard deviations are 2.60 and 2.88. In this case the estimated AUC is

$$\Phi\left[\frac{E\log X - E\log Y}{\{\mathrm{var}(\log X) + \mathrm{var}(\log Y)\}^{1/2}}\right] = 0.57.$$

Now, we test for $AUC = 0.5$ under the ROC curve of IL8 (i.e. no discriminatory ability of the biomarker). This is a particular case of the testing procedure considered in Section 2. Specifically, since the $AUC = 0.5$ iff $E\log X = E\log Y$, the test statistic has the form

$$2\log z = 2\left[\sup_{\mu_x,\mu_y,\sigma_x,\sigma_y}\{l(\log X, n_x, k_x; \mu_x, \sigma_x) + l(\log Y, n_y, k_y; \mu_y, \sigma_y)\}\right.$$
$$\left. - \sup_{\mu_x,\sigma_x,\sigma_y}\{l(\log X, n_x, k_x; \mu_x, \sigma_x) + l(\log Y, n_y, k_y; \mu_x, \sigma_y)\}\right],$$

where

$$l(z,n,k;\mu,\sigma) = -(n-k)\ln(\sigma) - \Sigma_{i:z_i>\log d}\frac{(z_i-\mu)^2}{2\sigma^2} + k\ln\Phi\left(\frac{\log d-\mu}{\sigma}\right),$$
$$n_x = 189,\ n_y = 180,\ k_x = 95,\ k_y = 108,$$

and the function $\exp(l)$ is proportional to the likelihood based on censored data. For details regarding this maximum likelihood function see Vexler *et al.* (2006). The value of the test-statistic is computed to be 2.97. Since the value of $z_{0.05}$ corresponding to $Pr_{H_0}(2\log z > z_{0.05}) \simeq 0.05$ (from χ_1^2 distribution) is 3.84, we do not reject H_0. Therefore we conclude that the discriminatory ability of biomarker IL8 is not significant.

4.2 *Evaluating Biomarkers for Coronary Heart Disease*

For this example, we compare the diagnostic accuracy of two biomarkers, cholesterol and hdl-cholesterol. To normalize the data, we log-transform the values of both biomarkers. It is obvious, from a biological standpoint that the levels of cholesterol and hdl-cholesterol are correlated. Denote by X_1, Y_1 the log-transformed values of cholesterol for the cases and controls, respectively, and similarly, X_2, Y_2 the log-transformed hdl-cholesterol levels from cases and controls. The estimated means of X_1, X_2, Y_1, and Y_2 are 5.63, 4.15, 5.47, 4.13, and the estimated standard deviations are 0.18, 0.24, 0.30, 0.25, respectively. The estimators of the correlation between X_1 and X_2, as well as between Y_1 and Y_2 are $\hat{\rho}_x = 0.06$ and $\hat{\rho}_y = 0.04$, respectively. Figure 2 introduces empirical histograms of X_1, X_2, Y_1, and Y_2.

[Figure 2 about here.]

Assume that the values of the log-transformed biomarkers are normally distributed. Simulation studies was conducted for each of the $d = 0$, 3, 3.25, 3.5, 3.75, 4, and 4.25. Table 5 presents estimators of the correlated AUCs and p-values obtained based on values of the test-statistic z for different d, (theoretically $2\log z^{(LOD)}$ is approximately χ_1^2-distributed). (Note that situation $d = 0$ corresponds to no LOD effect.)

[Table 5 about here.]

From Table 5, for any selected value of d, the null hypothesis H_0 is rejected with $p - values$ increasing as d increases.

Although standard `SPSS` output gives the asymptotic 95% confidence interval (CI) of AUC_1 as (0.628,0.708) and of AUC_2 as (0.481,0.585), in the simple case where $d = 0$, we can not conclude that $H_0 : AUC_1 = AUC_2$ is rejected because the estimators of AUC_1 and AUC_2 are correlated. We utilize a method proposed by Wieand *et al.* (1989, p. 587). Following these authors, we have, if biomarkers' values are normally distributed, the test-statistic

$$z_{CI} = \frac{\hat{\delta}_1 - \hat{\delta}_2}{\{(n+m)\text{var}(T)\}^{1/2}} \sim N(0,1),\ n, m \to \infty,$$

where

$$\begin{aligned}
\hat{\delta}_k &= \frac{\hat{\mu}_{x_k} - \hat{\mu}_{y_k}}{(\hat{\sigma}^2_{x_k} + \hat{\sigma}^2_{y_k})^{1/2}},\ \hat{\mu}_{x_k} = \frac{\Sigma_{i=1}^{n} x_{k_i}}{n},\ \hat{\mu}_{y_k} = \frac{\Sigma_{j=1}^{m} y_{k_j}}{m}, \\
\hat{\sigma}^2_{x_k} &= \frac{\Sigma_{i=1}^{n}(x_{k_i} - \hat{\mu}_{x_k})^2}{n-1},\ \hat{\sigma}^2_{y_k} = \frac{\Sigma_{j=1}^{m}(y_{k_j} - \hat{\mu}_{y_k})^2}{m-1}, \\
\text{var}((n+m)^{1/2}T) &= \sigma_{11} - 2\sigma_{12} + \sigma_{22}, \\
(n+m)^{-1}\sigma_{ii} &= \sigma_i^{-2}(n^{-1}\sigma^2_{x_i} + m^{-1}\sigma^2_{y_i}) + \frac{1}{2}\delta_i^2\sigma_i^{-4}\{(n-1)^{-1}\sigma^4_{x_i} \\
&\quad + (m-1)^{-1}\sigma^4_{y_i}\}, \\
(n+m)^{-1}\sigma_{12} &= (\sigma_1\sigma_2)^{-1}(n^{-1}C_x + m^{-1}C_y) + \frac{1}{2}\delta_1\delta_2(\sigma_1\sigma_2)^{-2}\{(n-1)^{-1}C_x^2 \\
&\quad + (m-1)^{-1}C_y^2\}, \\
\sigma_i^2 &= \sigma^2_{x_i} + \sigma^2_{y_i},\ C_x = \rho_x\sigma_{x_1}\sigma_{x_2},\ C_y = \rho_y\sigma_{y_1}\sigma_{y_2}, \\
\text{and}\ \ \delta_k &= \frac{\mu_{x_k} - \mu_{y_k}}{(\sigma^2_{x_k} + \sigma^2_{y_k})^{1/2}},\ k = 1, 2.
\end{aligned}$$

Thus, since the z_{CI} calculated from the data is 4.71, the $p-value$ of the test $\mid z_{CI} \mid > z_\alpha$ is 0.0021, whereas our proposed method has p-value 10.20×10^{-7}; see Table 5 ($d = 0$).

5. Discussion

In the present article, we have shown that the maximum likelihood ratio approach serves as a method of testing for the hypothesis regarding the comparison of AUCs. Such an approach yields a powerful test with characteristics that can be obtained by the well-established maximum likelihood theory. We used real data examples to illustrate how easily MLRT method can be carried out in order to compare two biomarkers and to determine whether a biomarker has discriminatory ability.

The paper assumes normal distributions for the values of the biomarkers when LOD is present. However, the proposed approach can be extended to other commonly used distributions, e.g., gamma, lognormal, etc. Similarly, we can perform hypothesis testing for AUCs based on right, double censored, or truncated data. We have focused on comparing paired correlated areas, but the proposed method can be adapted to multivariate cases as well.

Our paper presented a method dealing with data subject to LOD with broad validity under a reasonable set of assumptions. Sensitivity analysis, though beyond the scope of the present paper, is important to assess these distributional assumptions. This topic can be discussed in a general context of missing data analysis; see Molenberghs and Kenward (2007). However, one must bear in mind that data below LOD are informative missing, in the sense that they are unobservable only if the actual values are below the detection limit.

We briefly investigated several imputation methods that are commonly applied among epidemiologists in dealing with LOD data. These methods, however, are not statistically justified and should not be confused with the popular method of multiple imputation (e.g. Rubin, 1987) in the missing data analysis literature. The use of multiple imputation in the analysis of LOD data deserves further investigation.

Note that nonparametric distribution function estimation based on censored data can be obtained and hence Kolmogorov-Smirnov (or Shapiro-Wilk) -type tests for correctness of parametric assumptions can be evaluated (e.g., Verrill and Johnson, 1988). In the context of the ROC curves and Box-Cox power transformation models based on data subject to LOD, we will address nonparametric and semi parametric methods in a subsequent article.

The proposed approach preserves the efficiency of the MLRT when applied to testing for biomarkers' diagnostic accuracy subject to the limit of detection. When an additive measurement error is in effect, the appropriate maximum likelihood approach can also be utilized following method similar to that of Section 2.

Acknowledgements

This research was supported by the Intramural Research Program of the National Institute of Child Health and Human Development, National Institutes of Health. The opinions expressed are those of the authors and not necessarily of the National Institutes of Health. The authors would like to thank Margaret Hillier for providing the intrauterine inflammation data. We are grateful to the co-editor, associate editor, and referee for their helpful comments that clearly improved this paper.

References

Bamber, D. (1975). The Area above the Ordinal Dominance Graph and the Area below the Receiver Operating Characteristic Graph. *Journal of Mathematical Psychology* **12**, 387–415.

Choi, S., Hall, W. J. and Schick, A. (1996). Asymptotically Uniformly Most Powerfull Tests in Parametric and Semiparametric Models. *Ann. Stat.* **24**, 841-861.

Gupta, A. K. (1952). Estimation of the Mean and Standard Deviation of a Normal Population from a Censored Sample. *Biometrika* **39**, 260–273.

Lehmann, E. L. (1997). *Testing Statistical Hypotheses.* New York: John Wiley and Sons, 2nd edition.

Louis, G. M., Weiner, J. M., Whitecomb, B. W., Sperrazza, R., Schisterman, E. F., Lobdell, D. T., Crickard, K., Greizerstein, H., and Kostyniak, P. J. (2005). Environmental PCB exposure and risk of endometriosis. *Human Reproduction* **20**, 279–285.

Lubin, J. H., Colt, J. S., Camann, D., Davis, S., Cerhan, J. R., Severson, R. K., Bernstein, L., and Hartge, P. (2004). Epidemiological Evaluation of Measurement Data in the Presence of Detection Limits. *Environmental Health Perspectives* **112**, 1691–1696.

Lyles, R. H., Williams, J. K., and Chuachoowong, R. (2001). Correlating Two Viral Load Assays with Known Detection Limits. *Biometrics* **57**, 1238–1244.

Lynn, H.S. (2001). Maximum likelihood inference for left-censored HIV RNA data. *Statistics in Medicine* **20**, 33–45.

Molenberghs, G. and Kenward, M. G. (2007). *Missing Data in Clinical Studies.* Chichester: Wiley.

Molodianovitch, K., Faraggi, D., and Reiser, B. (2006). Comparing the Areas Under Two Correlated ROC Curves: Parametric and Non-Parametric

Approaches. *Biometrical Journal* **48**, 745–757.

Mumford, S. L., Schisterman, E. F., Vexler, A., and Liu, A. (2006). Pooling Biospecimens and Limits of Detection: Effects on ROC Curve Analysis. *Biostatistics* **7**, 585-598.

Perkins, N. J., Schisterman, E. F. and Vexler, A. (2006). Receiver operating characteristic curve inference from a sample with a Limit of Detection. *American Journal of Epidemiology*, in press.

Reiser, B. and Faraggi, D., (1997). Confidence Intervals for the Generalized ROC Criterion. *Biometrics* **53**, 644-652.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys.* New York: Wiley.

Schisterman, E. F., Vexler, A., Whitcomb, B. W., and Liu, A. (2006). The Limitations due to Exposure Detection Limits for Regression Models. *American Journal of Epidemiology* **163**, 374-383.

Shapiro, D. (1999). The interpretation of diagnositc tests. *Statistical Methods in Medical Research* **8**, 113–134.

Verrill, S., Johnson, R. A. (1988). Tables and Large-Sample Distribution Theory for Censored-Data Correlation Statistics for Testing Normality. *Journal of The American Statistical Association* **83**, 1192–1197.

Vexler, A., Liu, A., and Schisterman, E. F. (2006). Efficient Design and Analysis of Biospecimens with Measurements Subject to Detection Limit. *Biometrical Journal* **48**, 780–791.

Wieand, S., Gail, M. H., James, B. R., and James, K. L. (1989). A family of nonparametric statistics for comparing diagnostic markers with paired or unpaired data. *Biometrika* **76**, 585–92

Appendix A

The Likelihood Functions Based on Data Subject to LOD

The likelihood function based on $X' = \{x'_{1i}, x'_{2i}\}_{i=1}^{n}$ has the form of

$$L\left(X'; \Theta_X^{H_1}\right) = \left(\prod_{i=1}^{n_1} t_{i1}^{(x)}\right) \times \left(\prod_{i=n_1+1}^{n_2} t_{i2}^{(x)}\right) \times \left(\prod_{i=n_2+1}^{n_3} t_{i3}^{(x)}\right) \times (t_4^{(x)})^{n_4},$$

where

$$\begin{aligned} t_{i1}^{(x)} &= (2\pi\sigma_{x_1}\sigma_{x_2|x_1})^{-1} \exp\left[-0.5\left\{\frac{(x'_{2i} - \mu_{x_2|x_{1i}})^2}{\sigma_{x_2|x_1}^2} + \frac{(x'_{1i} - \mu_{x_1})^2}{\sigma_{x_1}^2}\right\}\right], \\ t_{i2}^{(x)} &= (2\pi\sigma_{x_1}^2)^{-1/2} \exp\left\{-0.5\frac{(x'_{1i} - \mu_{x_1})^2}{\sigma_{x_1}^2}\right\} \Phi\left(\frac{d - \mu_{x_2|x_{1i}}}{\sigma_{x_2|x_1}}\right), \end{aligned}$$

$$\begin{aligned} t_{i3}^{(x)} &= (2\pi\sigma_{x_2}^2)^{-1/2} \exp\left\{-0.5\frac{(x'_{2i} - \mu_{x_2})^2}{\sigma_{x_2}^2}\right\} \Phi\left(\frac{d - \mu_{x_1|x_{2i}}}{\sigma_{x_1|x_2}}\right), \\ t_4^{(x)} &= \int_{-\infty}^{d} \Phi\left[\frac{d - \left\{\mu_{x_1} + \frac{\rho_x\sigma_{x_1}(x_2 - \mu_{x_2})}{\sigma_{x_2}}\right\}}{\sigma_{x_1}\sqrt{1-\rho_x^2}}\right] (2\pi\sigma_{x_2}^2)^{-1/2} \\ &\quad \times \exp\left\{-0.5\frac{(x_2 - \mu_{x_2})^2}{\sigma_{x_2}^2}\right\} dx_2, \end{aligned}$$

$$\begin{aligned} \mu_{x_2|x_{1i}} &= \mu_{x_2} + (\rho_x\sigma_{x_2}/\sigma_{x_1})(x'_{1i} - \mu_{x_1}), \quad \sigma_{x_2|x_1}^2 = \sigma_{x_2}^2(1-\rho_x^2), \\ \mu_{x_1|x_{2i}} &= \mu_{x_1} + (\rho_x\sigma_{x_1}/\sigma_{x_2})(x'_{2i} - \mu_{x_2}), \quad \sigma_{x_1|x_2}^2 = \sigma_{x_1}^2(1-\rho_x^2), \end{aligned}$$

and n_1, n_2, n_3, n_4 $(n_1+n_2+n_3+n_4 = n)$ are the numbers of events $\{x'_{1i} \neq NA, x'_{2i} \neq NA\}_{i=1}^{n}$, $\{x'_{1i} \neq NA, x'_{2i} = NA\}_{i=1}^{n}$, $\{x'_{1i} = NA, x'_{2i} \neq NA\}_{i=1}^{n}$, and $\{x'_{1i} = NA, x'_{2i} = NA\}_{i=1}^{n}$, respectively. Here, the term $t_{i1}^{(x)}$ corresponds to situations where x_{1i} and x_{2i} are observed completely; $t_{i2}^{(x)}$ relates to situations where x_{1i} is observed, while x_{2i} is below the detection limit d and thus x'_{2i} is NA (the opposite case where x_{1i} is unobserved while x_{2i} is available, matches $t_{i3}^{(x)}$). Finally, when both x_1 and x_2 are not observed numerically, we have $t_4^{(x)}$ (i.e. $t_4^{(x)}$ represents the probability of both x_1' and x_2' to be NA).

The likelihood function based on $Y' = \{y_{1j}', y_{2j}'\}_{j=1}^{m}$ is defined in a similar manner:

$$\begin{aligned} L\left(Y'; \Theta_Y^{H_1}\right) &= \left(\prod_{j=1}^{m_1} t_{j1}^{(y)}\right) \times \left(\prod_{j=m_1+1}^{m_2} t_{j2}^{(y)}\right) \times \left(\prod_{j=m_2+1}^{m_3} t_{j3}^{(y)}\right) \times (t_4^{(y)})^{m_4}, \\ & \Sigma_{r=1}^{4} m_r = m. \end{aligned}$$

Appendix B

Transformed Normal Approach

We denote the function

$$T(u, \lambda) = \begin{cases} \frac{u^{\lambda}-1}{\lambda}, & \lambda \neq 0, \\ \log(u), & \lambda = 0 \end{cases}$$

and extend the likelihoods $L\left(X'; \Theta_X^{H_1}\right)$ and $L\left(Y'; \Theta_Y^{H_1}\right)$ by Appendix A to the forms of

$$\begin{aligned} L\left(X'; \Theta_X^{H_1}, \lambda_1, \lambda_2\right) &= \left(\prod_{i=1}^{n_1} \widetilde{t}_{i1}^{(x)}\right) \times \left(\prod_{i=n_1+1}^{n_2} \widetilde{t}_{i2}^{(x)}\right) \times \left(\prod_{i=n_2+1}^{n_3} \widetilde{t}_{i3}^{(x)}\right) \times (\widetilde{t}_4^{(x)})^{n_4}, \\ L\left(Y'; \Theta_Y^{H_1}, \lambda_1, \lambda_2\right) &= \left(\prod_{i=1}^{m_1} \widetilde{t}_{i1}^{(y)}\right) \times \left(\prod_{i=m_1+1}^{m_2} \widetilde{t}_{i2}^{(y)}\right) \times \left(\prod_{i=m_2+1}^{m_3} \widetilde{t}_{i3}^{(y)}\right) \times (\widetilde{t}_4^{(y)})^{m_4}, \end{aligned}$$

where for $z = x, y$

$$\begin{aligned} \widetilde{t}_{i1}^{(z)} &= (2\pi\sigma_{z_1}\sigma_{z_2|z_1})^{-1} \\ & \quad \times \exp\left[-0.5\left\{\frac{(T(z_{2i}', \lambda_2) - \mu_{z_2|z_{1i}})^2}{\sigma_{z_2|z_1}^2} + \frac{(T(z_{1i}', \lambda_1) - \mu_{z_1})^2}{\sigma_{z_1}^2}\right\}\right], \\ \widetilde{t}_{i2}^{(z)} &= (2\pi\sigma_{z_1}^2)^{-1/2} \exp\left\{-0.5\frac{(T(z_{1i}', \lambda_1) - \mu_{z_1})^2}{\sigma_{z_1}^2}\right\} \Phi\left(\frac{T(d, \lambda_2) - \mu_{z_2|z_{1i}}}{\sigma_{z_2|z_1}}\right), \end{aligned}$$

$$
\begin{aligned}
\widetilde{t}_{i3}^{(z)} &= (2\pi\sigma_{z_2}^2)^{-1/2}\exp\left\{-0.5\frac{(T(z'_{2i},\lambda_2)-\mu_{z_2})^2}{\sigma_{z_2}^2}\right\}\Phi\left(\frac{T(d,\lambda_1)-\mu_{z_1|z_{2i}}}{\sigma_{z_1|z_2}}\right),\\
\widetilde{t}_{4}^{(z)} &= \int_{-\infty}^{T(d,\lambda_2)}\Phi\left[\frac{T(d,\lambda_1)-\left\{\mu_{z_1}+\frac{\rho_z\sigma_{z_1}(z_2-\mu_{z_2})}{\sigma_{z_2}}\right\}}{\sigma_{z_1}\sqrt{1-\rho_z^2}}\right](2\pi\sigma_{z_2}^2)^{-1/2}\\
&\quad\times\exp\left\{-0.5\frac{(z_2-\mu_{z_2})^2}{\sigma_{z_2}^2}\right\}dz_2,
\end{aligned}
$$

and it is assumed that $(T(z_1,\lambda_1),T(z_2,\lambda_2))$ are jointly normally distributed. Thus, in this case, the MLR test-statistic is

$$
\frac{\sup_{\bar{\mu}_{x_1},\bar{\mu}_{x_2},\bar{\mu}_{y_1},\bar{\mu}_{y_2},\bar{\sigma}_{x_1}^2,\bar{\sigma}_{x_2}^2,\bar{\sigma}_{y_1}^2,\bar{\sigma}_{y_2}^2,\bar{\rho}_x,\bar{\rho}_y,\lambda_1,\lambda_2} L\left(X';\overline{\Theta}_X^{H_1},\lambda_1,\lambda_2\right)L\left(Y';\overline{\Theta}_Y^{H_1},\lambda_1,\lambda_2\right)}{\sup_{\bar{\mu}_{x_2},\bar{\mu}_{y_1},\bar{\mu}_{y_2},\bar{\sigma}_{x_1}^2,\bar{\sigma}_{x_2}^2,\bar{\sigma}_{y_1}^2,\bar{\sigma}_{y_2}^2,\bar{\rho}_x,\bar{\rho}_y,\lambda_1,\lambda_2} L\left(X';\overline{\Theta}_X^{H_0},\lambda_1,\lambda_2\right)L\left(Y';\overline{\Theta}_Y^{H_0},\lambda_1,\lambda_2\right)}.
$$

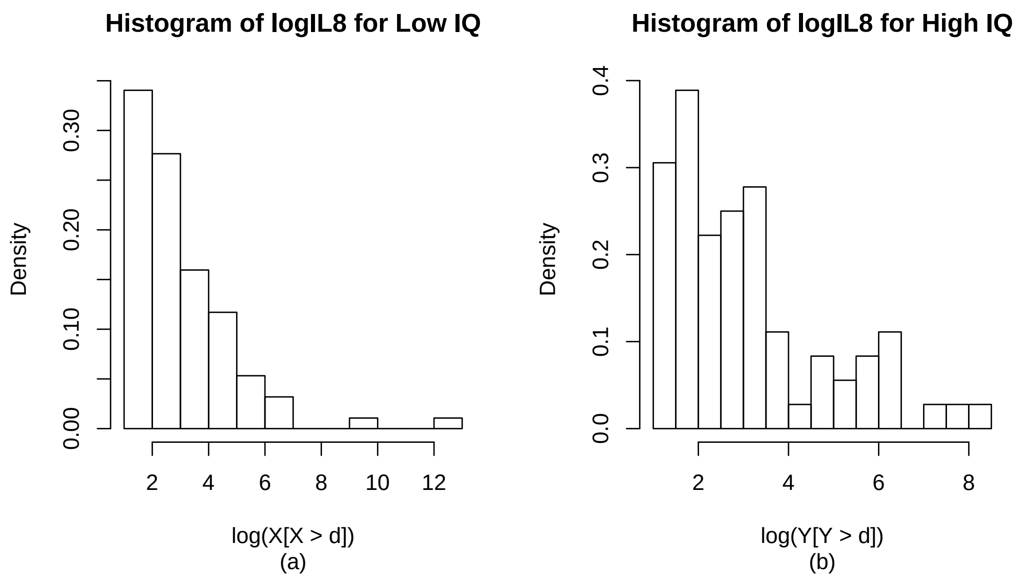

Figure 1. Histograms of the log-transformed biomarker of interest corresponding to low (a) and high (b) levels of IQ.

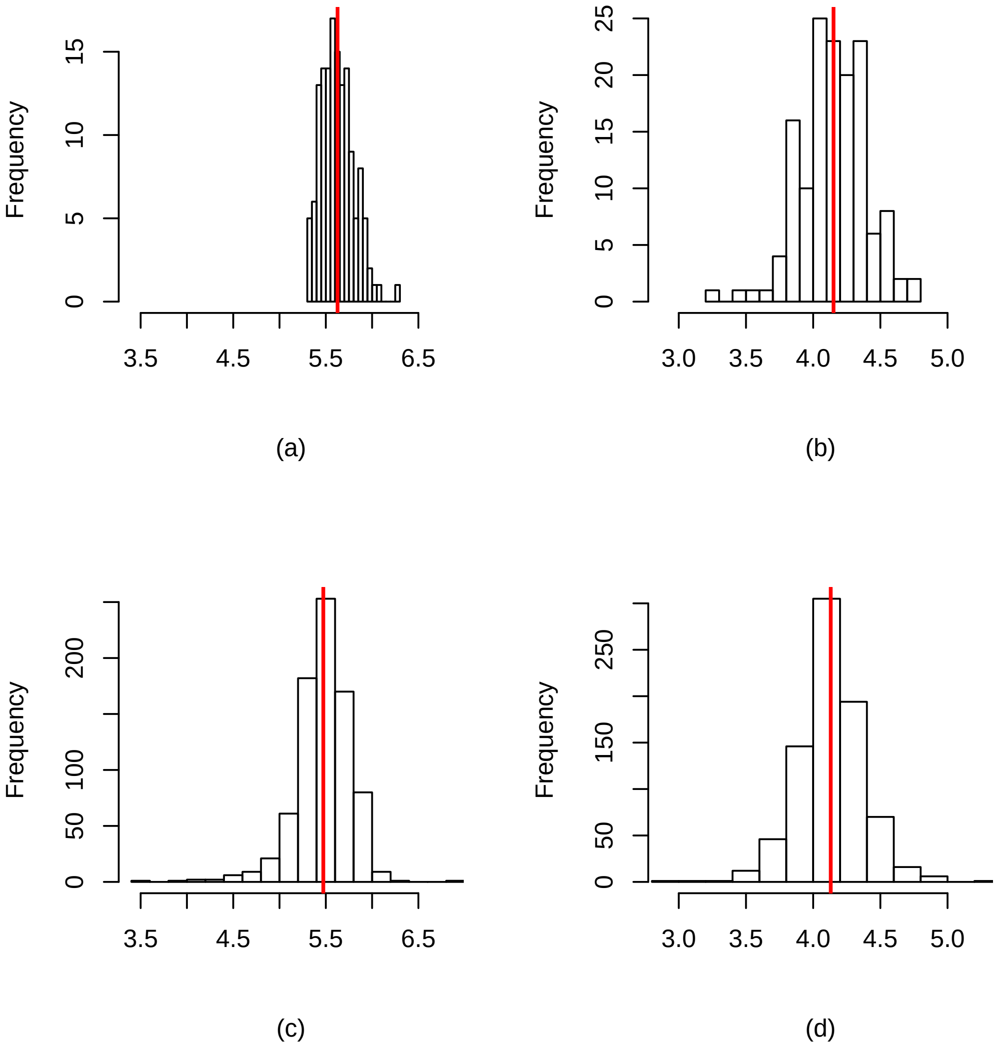

Figure 2. Histograms of the log-transformed biomarkers of interest corresponding to cholesterol cases (a), hdl-cholesterol cases (b), cholesterol controls (c), and hdl-cholesterol controls (d). The vertical bold lines correspond to average values of the biomarkers.

Table 1

Monte Carlo results for the significance levels of the proposed test.
$F(u) = Pr(2 \log z^{(LOD)} > u)$ *and* $F(3.84) \simeq 0.05$, $F(6.63) \simeq 0.01$.

$n = m$	d	$F(3.84)$	$F(6.63)$	P($x_1 < d$)	P($y_1 < d$)	P($x_2 < d$)	P($y_2 < d$)
150	-3	0.0504	0.0098	9.5×10^{-6}	6.22×10^{-16}	0.0007	0.0003
	-1	0.0510	0.0100	0.0115	3.17×10^{-5}	0.0606	0.0728
	-0.5	0.0535	0.0114	0.0380	0.0013	0.1271	0.1660
	0	0.0573	0.0146	0.1013	0.0228	0.2325	0.3138
	0.5	0.0601	0.0153	0.2194	0.1587	0.3740	0.5000
	0.75	0.0634	0.0241	0.3000	0.3085	0.4537	0.5958
30	-3	0.0510	0.0110				
	-0.5	0.0507	0.0120				
	0	0.0593	0.0150				

Table 2

Monte Carlo results for the power of the test. $F = Pr(2\log z^{(LOD)} > 3.84)$.

d	AUC_1	AUC_2	F	P($x_1 < d$)	P($y_1 < d$)	P($x_2 < d$)	P($y_2 < d$)
-3	0.5	0.6	0.8394	3.2×10^{-5}	6.2×10^{-16}	1.1×10^{-3}	5.5×10^{-4}
-1	0.5	0.6	0.8284	2.3×10^{-2}	3.2×10^{-5}	8.2×10^{-2}	9.9×10^{-2}
0	0.5	0.6	0.8374	0.16	2.3×10^{-2}	0.28	0.38
0.75	0.5	0.6	0.7372	0.40	0.31	0.52	0.67
-3	0.6	0.9	0.9995	8.5×10^{-6}	6.2×10^{-16}	5.5×10^{-4}	0.12
-1	0.6	0.9	0.9985	1.1×10^{-2}	3.2×10^{-5}	0.07	0.66
0	0.6	0.9	0.9973	0.08	0.02	0.28	0.89

Table 3

The Monte Carlo Type I error of the proposed test given the uncorrected distributional assumption.

df	d	$\hat{P}r(2\log z^{(LOD)} > 3.84)$
15	-3	0.0506
10	-3	0.0524
5	-3	0.0518
15	-1	0.0590
10	-1	0.0610
5	-1	0.0651
15	0	0.0958
10	0	0.1507
5	0	0.3288

Table 4

The Monte Carlo Type I error of the test based on confidence intervals when the imputation method is applied and the expected significance level is 0.05.

d	$Imp = d/2$	$Imp = d/\sqrt{2}$
-3	0.0549	0.0546
0	0.0689	0.0687
0.5	0.1098	0.1455
0.75	0.1539	0.2273

Table 5

Estimation of the AUCs and values of the test-statistic for different d. N_{X_k}, N_{Y_k} are numbers of events $\{X_k < d\}$ and $\{Y_k < d\}$, respectively $(k = 1, 2)$.

d	N_{X_1}	N_{X_2}	N_{Y_1}	N_{Y_2}	$A\hat{U}C_1$	$A\hat{U}C_2$	$2\log z^{(LOD)}$	$p-value$
0.00	0	0	0	0	0.671	0.524	24.291	10.20×10^{-7}
3.00	0	0	0	1	0.671	0.524	24.049	9.39×10^{-7}
3.25	0	0	0	2	0.671	0.524	22.156	2.51×10^{-6}
3.50	0	2	1	8	0.671	0.524	21.969	2.77×10^{-6}
3.75	0	8	1	40	0.672	0.525	22.570	2.03×10^{-6}
4.00	0	34	2	207	0.672	0.530	22.941	1.67×10^{-6}
4.25	0	88	5	553	0.673	0.583	21.862	2.93×10^{-6}