Combining biomarkers to improve diagnostic accuracy using the overlap coefficient

Tahani Coolen-Maturi*

Department of Mathematical Sciences, Durham University, Durham DH1 3LE, UK.

Abstract

Measuring the accuracy of diagnostic tests is crucial in many application areas, including medicine, machine learning, and credit scoring. In practice, multiple diagnostic tests or biomarkers are combined to improve diagnostic accuracy. The area under the receiver operating characteristic curve (AUC) is a common measure of diagnostic test performance and can be used as an objective function to maximise when combining multiple biomarkers. Another useful measure is the overlap coefficient, which quantifies the similarity between two independent distributions by their overlapping area. The smaller the overlapping area, the better the biomarker is at discrimination. The aim of this paper is to combine biomarkers to improve diagnostic accuracy by minimising the overlap coefficient. We approach this parametrically and non-parametrically using Kernel-based methods. We also present a probabilistic interpretation of the overlap coefficient, which gives more insight into this measure. The proposed methods are evaluated through a simulation study and illustrated via examples.

Keywords: Diagnostic accuracy; combining biomarkers; overlap coefficient; ROC; AUC

1 Introduction

Measuring the accuracy of diagnostic tests is essential in various fields, such as medicine, machine learning, and credit scoring. In practical situations, a single diagnostic test may not be sufficient to make a useful decision. Hence, multiple diagnostic tests or biomarkers are often combined to improve diagnostic accuracy [1]. The area under the receiver operating characteristic curve (AUC) is a common measure of diagnostic test performance. It can be used as an objective function to maximise when combining multiple biomarkers.

*Email: tahani.maturi@durham.ac.uk (Corresponding author)

Another useful measure, yet less popular, is the overlap coefficient, denoted by OVL hereafter. It quantifies the similarity between two independent distributions by their overlapping area. The smaller the overlapping area, the better the biomarker is at discrimination. The OVL coefficient is highly intuitive and has an easy visual representation. Inferential methods, both parametric and non-parametric, have been developed in the literature, see e.g. [2, 3, 4]. Unlike the AUC, the OVL is able to capture shape differences For more details about OVL versus AUC, we refer the reader to Wang and Tian [5]. The aim of this paper is to combine biomarkers to improve diagnostic accuracy by minimising the overlap coefficient. We approach this problem both parametrically and non-parametrically using Kernel-based methods. We also present a probabilistic interpretation of the overlap coefficient, which provides more insight into this measure. We will compare the combination of biomarkers using both measures, AUC and OVL, through simulation studies and examples.

The rest of the paper is organised as follows: Section 2 introduces the overlap coefficient (OVL) for one biomarker, with the new probabilistic interpretation given in Section 2.1. The main results are given in Section 3, where multiple biomarkers are combined to improve the diagnostic accuracy by minimising the overlap coefficient. The proposed methods are assessed through a simulation study in Section 4. An example is provided to illustrate the proposed method in Section 5. The paper ends with some concluding remarks in Section 6. The R code used for implementation is available from the author on request.

2 OVL for one biomarker

Suppose that X is a continuous random quantity of a diagnostic test result and that larger values of X are considered more indicative of disease. Here, X_0 and X_1 are used to denote test results for the non-disease (control) and disease (cases) groups, respectively. A useful summary is the area under the ROC curve, $\text{AUC} = \int_0^1 \text{ROC}(t)\, dt$. The AUC measures the overall performance of the diagnostic test. Higher AUC values indicate more accurate tests, with AUC = 1 for perfect or ideal tests and AUC = 0.5 for uninformative tests. The AUC is equal to the probability that the test results from a randomly selected pair of diseased and non-diseased subjects are correctly ordered, i.e., $\text{AUC} = P(X_1 > X_0)$ [6]. Thus, the AUC measures the test's ability to correctly classify a randomly selected individual as being from either the disease group or the non-disease group. The Gini coefficient is a simple conversion from AUC using the formula $\text{Gini} = 2\text{AUC} - 1$. Both measures are mathematically equivalent and interchangeable. In some contexts, such as clinical settings, AUC is often preferred due to its direct interpretability, where AUC = 0.5 indicates random performance. On the other hand, the Gini coefficient has a more intuitive scale, ranging from 0 to 1, compared to the AUC scale of 0.5 to 1. However, all the properties and limitations of AUC

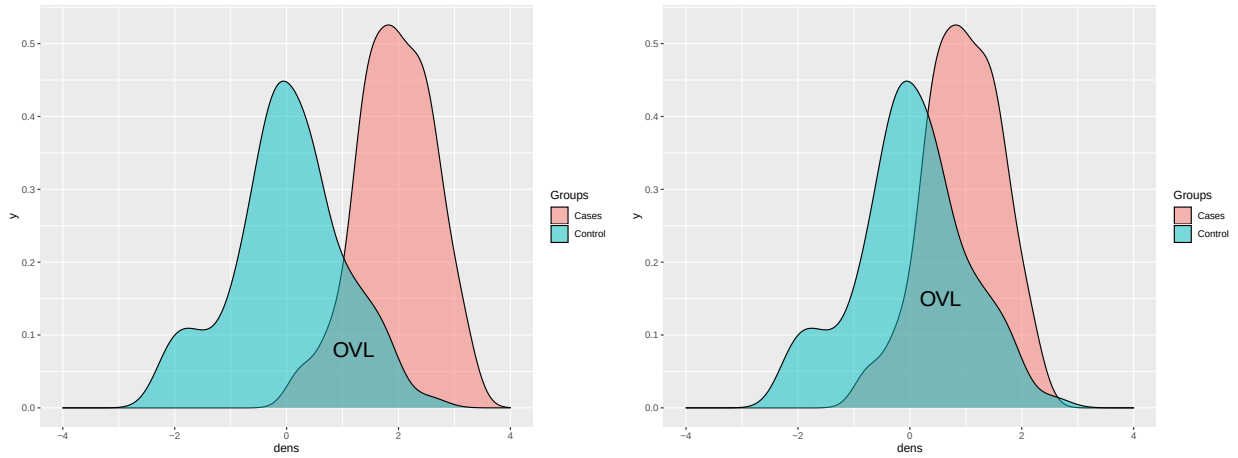

Figure 1: Low OVL (left) and high OVL (right)

apply to the Gini coefficient as well [7].

Another useful but less popular measure is the Overlap Coefficient (OVL), which measures the overlapping area of the densities of two distributions when plotted on the same axes, e.g., the two plots in Figure 1 are examples of low and high overlapping probability densities. The OVL takes a value between 0 and 1, where OVL = 0 if the two probability densities are disjoint and OVL = 1 if the two densities are the same. There are no clear cutoff values to describe the discrimination ability of this measure, but a rule of thumb has been suggested by Franco-Pereira et al. [3] as follows:

$$\begin{cases} \text{OVL} = 1 & \text{no differentiation;} \\ 0.75 < \text{OVL} < 1 & \text{poor differentiation;} \\ 0.55 < \text{OVL} < 0.75 & \text{good differentiation;} \\ 0.35 < \text{OVL} < 0.55 & \text{very good differentiation;} \\ \text{OVL} < 0.35 & \text{excellent differentiation.} \end{cases}$$

Formally, let f_{X_0} and f_{X_1} be the corresponding probability densities, for the non-disease (control) and disease (cases) groups, respectively. The overlap coefficient is the overlap area between the two densities, which is defined as

$$\text{OVL} = \int \min[f_{X_0}(x), f_{X_1}(x)]dx \tag{1}$$

One can estimate the OVL coefficient parametrically by relying on the normality assumption of both densities as follows: Let $X_{0,1}, X_{0,2}, \ldots, X_{0,n_0}$ and $X_{1,1}, X_{1,2}, \ldots, X_{1,n_1}$ denote two random samples of sizes n_0 and n_1 taken from two independent normal distributions with means μ_0 and μ_1 and standard deviations σ_0 and σ_1, respectively, thus OVL is defined as [3]

$$\text{OVL} = \Phi\left(\frac{d_0 - \mu_0}{\sigma_0}\right) - \Phi\left(\frac{d_0 - \mu_1}{\sigma_1}\right) + \Phi\left(\frac{d_1 - \mu_1}{\sigma_1}\right) - \Phi\left(\frac{d_1 - \mu_0}{\sigma_0}\right) + 1 \tag{2}$$

where, for $\sigma_0 < \sigma_1$, the intersection points d_0 and d_1 are given by

$$d_0 = \frac{(\mu_0\sigma_1^2 - \mu_1\sigma_0^2) - \sigma_0\sigma_1\sqrt{(\mu_0 - \mu_1)^2 + (\sigma_0^2 - \sigma_1^2)\log(\sigma_0^2/\sigma_1^2)}}{(\sigma_1^2 - \sigma_0^2)}$$

$$d_1 = \frac{(\mu_0\sigma_1^2 - \mu_1\sigma_0^2) + \sigma_0\sigma_1\sqrt{(\mu_0 - \mu_1)^2 + (\sigma_0^2 - \sigma_1^2)\log(\sigma_0^2/\sigma_1^2)}}{(\sigma_1^2 - \sigma_0^2)}$$

where Φ is the cumulative distribution function of the standard normal distribution, and the maximum likelihood estimates of the unknown parameters, μ_i and σ_i, $i = 0, 1$, can be used. We are also interested in the following special cases: when $\mu_0 = 0$ and $\sigma_0 = 1$, we have

$$\text{OVL} = \Phi(d_0) - \Phi\left(\frac{d_0 - \mu_1}{\sigma_1}\right) + \Phi\left(\frac{d_1 - \mu_1}{\sigma_1}\right) - \Phi(d_1) + 1$$

where, for $\sigma_1 > 1$, the intersection points d_0 and d_1 are given by

$$d_0 = \frac{-\mu_1 - \sigma_1\sqrt{\mu_1^2 + (\sigma_D^2 - 1)\log(\sigma_1^2)}}{(\sigma_1^2 - 1)}, \;\; d_1 = \frac{-\mu_1 + \sigma_1\sqrt{\mu_1^2 + (\sigma_D^2 - 1)\log(\sigma_1^2)}}{(\sigma_1^2 - 1)}$$

and when $\sigma_0 = \sigma_1 = \sigma$, this reduces to $\text{OVL} = 2\Phi(-|\mu_1 - \mu_0|/2\sigma)$.

The normality assumption can be quite restrictive, which may lead to inaccurate results when this assumption is violated, so a common way to extend the applicability of the binormal model is to apply a monotone transformation through the use of the Box-Cox transformation [3]. Alternatively, several Kernel-based approaches through the use of the Gaussian kernel using different bandwidths have been introduced and studied in the literature [4]. In this case, the densities in Equation (1) are replaced by appropriate kernel density estimators, e.g. the density estimator for $f_{X_0}(x)$ is given by

$$\hat{f}_{X_0}(x) = \frac{1}{n_0}\sum_{i=1}^{n_0}\frac{1}{h_0}K\left(\frac{x - X_{0,i}}{h_0}\right) \tag{3}$$

where K is a kernel function, and the bandwidth is given by

$$h_0 = (4/3)^{1/5}(n_0)^{-1/5}s_0, \quad \text{where } \; s_0 = \sqrt{\frac{1}{n_0 - 1}\sum_{i=1}^{n_0}\left(X_{0,i} - \sum_{j=1}^{n_0}\frac{X_{0,j}}{n_0}\right)^2}$$

We can define $\hat{f}_{X_1}(x)$ similarly, and we denote the resulting estimator by $\widehat{\text{OVL}}$ [3].

For the purpose of comparison, we define the empirical estimate for the overlap coefficient (OVL) as follows:

$$\widehat{\text{OVL}}_e = 1 - \max_x |\hat{F}_0(x) - \hat{F}_1(x)|$$

which is equivalent to the two-sample Kolmogorov-Smirnov test statistic [8], where $\hat{F}_0$ and $\hat{F}_1$ are the empirical distribution functions (edf) for the non-disease (control) and disease (cases) groups, respectively.

2.1 Interpretation for OVL

Suppose we have two groups or classes, A and B, and let us consider the natural (e.g. likelihood-based) classification rule, such that we assign an individual with observation x to A if $f_a(x) > f_b(x)$ and to B if $f_b(x) > f_a(x)$. Then we can divide the OVL into two areas (neglecting any x values for which $f_a(x) = f_b(x)$; if the set of such x values is not just one or a finite number of singletons, one may e.g. just assign randomly to A or B): OVL_A and OVL_B, where OVL_A is the area under $f_a(\cdot)$ for $\{x|f_a(x) < f_b(x)\}$ and OVL_B is the area under $f_b(\cdot)$ for $\{x|f_b(x) < f_a(x)\}$. So, OVL_A is the probability that an individual from group A is wrongly classified as belonging to group B, and OVL_B is the probability that an individual from group B is wrongly classified as belonging to group A. Now, let us consider a person on whom the diagnostic test is performed. It is unknown to which group this person belongs; assume he belongs to group A and to group B each with probability 0.5. Let WC be the event that this person is wrongly classified, then using, e.g. A to denote that the person belongs to group A,

$$\begin{aligned} P(WC) &= P(WC|A)P(A) + P(WC|B)P(B) \\ &= \frac{1}{2}\left[\int 1\{f_a(x) < f_b(x)\} f_a(x)dx + \int 1\{f_b(x) < f_a(x)\} f_b(x)dx\right] \\ &= \frac{1}{2}(\text{OVL}_A + \text{OVL}_B) \end{aligned}$$

So, $\text{OVL} = \text{OVL}_A + \text{OVL}_B = 2P(WC)$, so minimising OVL simply means minimising the probability that a person is wrongly classified, assuming he is equally likely to actually belong to groups A and B, and using the basic classification rule outlined above.

An extreme case is if $f_a(x) = f_b(x)$ for all x, so they have the same probability distribution; this would not work with the above classification rule but for any rule to classify as belonging to A if x is in some subset of the possible values, and to B otherwise, we would have $\text{OVL} = \text{OVL}_A + \text{OVL}_B = 1$ and hence $P(WC) = 1/2$, which is correct in that case. If the probability distributions of the two groups are completely separated, so no overlap, then we have $P(WC) = 0$, and hence $\text{OVL} = 0$.

3 Combining multiple biomarkers to minimise OVL

A practical question that often arises is how to effectively combine information from multiple diagnostic tests or biomarkers to accurately differentiate between diseased and non-diseased groups. In this section, our focus is on constructing linear combinations of diagnostic results, denoted as $Y = \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_p X_p$, where the objective is to determine the optimal values of the vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p)^T$ that minimise the overlap coefficient and, consequently, maximise the diagnostic accuracy.

The concept of linearly combining biomarkers to improve diagnostic accuracy by maximising the area under the ROC curve (AUC) has been explored by Lu [9] under the assumption of normality, while others proposed distribution-free approaches to relax this assumption, see, e.g. [1, 10, 11]. In this paper, we compare the proposed method with those based on AUC, with a slight difference in the approach of finding the optimal coefficients for the linear combination. Therefore, our objective is to find the optimal unit p-simplex vector $\alpha = (\alpha_1, \ldots, \alpha_p)^T$, such that the AUC is maximised or the OVL is minimised. A unit simplex vector is defined as a vector with non-negative values, where the entries sum up to one. The goal is to improve diagnostic accuracy by combining multiple diagnostic tests or biomarkers instead of relying on individual biomarkers alone, resulting in higher AUC or lower OVL values.

We start by introducing key notations, followed by an exploration of determining the optimal linear combination of biomarkers. Let us consider a set of p biomarkers, denoted as $X_1, X_2, \ldots, X_p$ or simply as $\boldsymbol{X} = (X_1, X_2, \ldots, X_p)^T$. We can define the linear combinations of the diagnostic results for the diseased and non-diseased groups as

$$Y_1 = \alpha_1 X_{1,1} + \alpha_2 X_{1,2} + \ldots + \alpha_p X_{1,p} = \boldsymbol{\alpha}^T \boldsymbol{X}_1$$
$$Y_0 = \alpha_1 X_{0,1} + \alpha_2 X_{0,2} + \ldots + \alpha_p X_{0,p} = \boldsymbol{\alpha}^T \boldsymbol{X}_0$$

where $\boldsymbol{X}_1 = (X_{1,1}, X_{1,2}, \ldots, X_{1,p})^T$ and $\boldsymbol{X}_0 = (X_{0,1}, X_{0,2} \ldots, X_{0,p})^T$.

Now let f_{Y_0} and f_{Y_1} be the corresponding probability densities, and the overlap area between these two densities is given by

$$\text{OVL}_Y = \int \min[f_{Y_0}(y), f_{Y_1}(y)]dy$$

The main questions that need to be addressed are how to estimate the densities f_{Y_0} and f_{Y_1}, and how to obtain the optimal values of $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_p)^T$ that minimise the overlap coefficient (OVL) for the combined test. We estimate the two densities, first parametrically, by assuming that the test results are generated from multivariate normal distributions and non-parametrically using Kernel-based methods.

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_p)^T$ be generated from multivariate normal distributions for n_1 cases and n_0

controls with means $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_0$ and variances $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_0$, respectively. Hence, the linear combinations of the diagnostic results for the diseased (cases) and non-diseased (control) groups are also normal, i.e. we have

$$Y_1 = \alpha_1 X_{1,1} + \alpha_2 X_{1,2} + \ldots + \alpha_p X_{1,p} = \boldsymbol{\alpha}^T \boldsymbol{X}_1 \sim N(\mu_{Y_1} = \boldsymbol{\alpha}^T \boldsymbol{\mu}_1, \sigma^2_{Y_1} = \boldsymbol{\alpha}^T \boldsymbol{\Sigma}_1 \boldsymbol{\alpha})$$
$$Y_0 = \alpha_1 X_{0,1} + \alpha_2 X_{0,2} + \ldots + \alpha_p X_{0,p} = \boldsymbol{\alpha}^T \boldsymbol{X}_0 \sim N(\mu_{Y_0} = \boldsymbol{\alpha}^T \boldsymbol{\mu}_0, \sigma^2_{Y_0} = \boldsymbol{\alpha}^T \boldsymbol{\Sigma}_0 \boldsymbol{\alpha})$$

Now Equation (2) can be applied to the combined scores, Y_1 and Y_0, by substituting the means and variances with $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_0$, $\boldsymbol{\Sigma}_1$, and $\boldsymbol{\Sigma}_0$. By utilising the maximum likelihood (ML) estimates of these means and variances, we can compute the corresponding intersection points, and hence find the parametric estimate of OVL_Y.

Again, the normality assumption can be quite restrictive, even more so in the multivariate case. To this end, one can use kernel-based approaches to estimate these probability densities. That is, the kernel-based estimates of the probability densities of the combined scores, Y_1 and Y_0, are given by

$$\hat{f}_{Y_1}(t) = \sum_{i=1}^{p} \alpha_i \hat{f}_{X_{1,i}}(x)$$
$$\hat{f}_{Y_0}(t) = \sum_{i=1}^{p} \alpha_i \hat{f}_{X_{0,i}}(x)$$

where $\hat{f}_{X_{1,i}}(x)$ and $\hat{f}_{X_{0,i}}(x)$, $i = 1, \ldots, p$, are the kernel density estimates calculated for each biomarker using Equation (3). These estimated densities are then substituted into Equation (1) to compute the kernel-based OVL_Y estimate.

To determine the optimal $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_p)^T$, we propose to find the values that minimise the OVL, where $\alpha_j \in [0, 1]$, $j = 1, 2, \ldots, p$ and $\sum_j \alpha_j = 1$. We will search for the optimal values of $\alpha_1, \ldots, \alpha_p$, where the OVL corresponding to the combined test $\sum_j \alpha_j X_j$ is evaluated for 101 equally spaced values for each $\alpha_j \in [0, 1]$, $j = 1, \ldots, p$ such that $\sum_j \alpha_j = 1$. We will use a similar approach for AUC, but instead of minimising the OVL, we will maximise it.

4 Simulation study

In this section, a simulation study is conducted to illustrate the proposed method for different scenarios. We have simulated two diagnostic test results (X_1, X_2) from the bivariate normal distribution for n_1 cases (disease) and n_0 controls (non-disease), with mean and variance-covariance matrix for the cases (disease)

and for the controls (non-disease), respectively,

$$\boldsymbol{m}_1 = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \boldsymbol{\Sigma}_1 = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}, \boldsymbol{m}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{\Sigma}_0 = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

without loss of generality we assume that $\mu_1 > \mu_2 > 0$, $\sigma_1^2 \geq 1$ and $\sigma_2^2 \geq 1$, and we considered $\rho \geq 0$, the correlation between X_1 and X_2, to be of most practical interest [1]. The area under the ROC curve for biomarker X_1 alone is equal to $\text{AUC}_1 = \Phi(\mu_1/\sqrt{1+\sigma_1^2})$, while the corresponding OVL measure for this biomarker can be obtained using the special cases formulas in Section 2. The same can be defined for biomarker X_2. The empirical estimate, $\widehat{\text{AUC}}_e$, and the kernel-based estimate $\widehat{\text{AUC}}_k$ are given by [12]

$$\widehat{\text{AUC}}_e = \frac{1}{n_1 n_0} \sum_{j=1}^{n_0} \sum_{i=1}^{n_1} \left[\mathbf{1}\left\{x_i^1 > x_j^0\right\} + \frac{1}{2}\mathbf{1}\left\{x_i^1 = x_j^0\right\} \right]$$

$$\widehat{\text{AUC}}_k = \frac{1}{n_0 n_1} \sum_{j=1}^{n_0} \sum_{i=1}^{n_1} \Phi\left(\frac{x_i^1 - x_j^0}{\sqrt{h_0^2 + h_1^2}} \right)$$

with the bandwidths $h_0 = 0.9 \times \min(s_0, \text{IQR}_0/1.34) \times n_0^{-1/5}$ and $h_1 = 0.9 \times \min(s_1, \text{IQR}_1/1.34) \times n_1^{-1/5}$, where s_0 and s_1 are the sample variances and where IQR_0 and IQR_1 are the sample interquartile ranges.

The results of the simulation study are based on 1000 simulations for $n_0 = n_1 = 50, 100$, $\rho = 0, 0.5, 0.75$, and for different values of μ_1, μ_2, σ_1, and σ_2. To facilitate comparison, we report the values of the Gini index, $G = 2\text{AUC} - 1$, instead of AUC, and $V = 1 - \text{OVL}$ instead of OVL. This allows the values to range from 0 to 1, where 0 represents no separation, and 1 represents complete separation of the disease and non-disease groups. The goal is to determine the best values for α_1 and α_2 that maximise both G and V. Here, $\widehat{G_b}$ and $\widehat{V_b}$ are the parametric estimates, $\widehat{G_e}$ and $\widehat{V_e}$ are the empirical estimates, and $\widehat{G_k}$ and $\widehat{V_k}$ are the nonparametric estimates. We consider the means of biomarker measurements corresponding to $\text{AUC} = 0.6, 0.7, 0.8, 0.9$ ($G = 0.2, 0.4, 0.6, 0.8$) as $0.358, 0.742, 1.190, 1.812$, respectively. The corresponding values of OVL are 0.8578, 0.7108, 0.5518, 0.3648, respectively ($V = 0.1422, 0.2892, 0.4482, 0.6352$).

We will distinguish between three scenarios depending on whether the disease and non-disease groups have equal variances. In the first scenario, we consider the case of equal variances, where $\sigma_1^2 = \sigma_2^2 = 1$, indicating that all variances are equal for both biomarkers and for the disease and non-disease groups. The second scenario is when $\sigma_1^2 > 1$ and $\sigma_2^2 = 1$ (i.e., the disease and non-disease groups have equal variances for biomarker 2 and different variances for biomarker 1). In the final scenario, we consider the case where the variances are unequal, that is, when $\sigma_1^2 > \sigma_2^2 > 1$ (i.e., the variance of the disease group of biomarker 1 is greater than the variance of the disease group in biomarker 2).

The simulation results, now detailed in Tables 4-9 in the Appendix, demonstrate an overall improvement in diagnostic accuracy when two biomarkers are combined. When the individual accuracy of the biomarkers is equal, they are assigned equal weights in the combined scores. However, if one biomarker exhibits higher diagnostic accuracy, it is given more weight, as expected. In cases where biomarkers are highly correlated, the weights tend to be heavily skewed towards the biomarker with the higher accuracy. Generally, the most significant improvements in diagnostic accuracy are observed when combining two independent biomarkers of equal accuracy.

5 Examples

In this section, we are providing two examples to demonstrate the proposed methods. In these examples, we have combined biomarkers using three different approaches: parametric, empirical, and nonparametric. The aim is to minimise the overlapping coefficient and improve the accuracy of the results. We have also compared the OVL results obtained from these methods with the AUC results.

Example 1 (**Pancreatic cancer data set).** The data set used in this example is from a study that included 90 pancreatic cancer patients and 51 control patients with pancreatitis [13]. Two serum markers, the cancer antigen CA125 and the carbohydrate antigen CA19-9, were measured. To make the results comparable, the marker values were transformed to a natural logarithmic scale and standardised; these are displayed in Figure 2. For simplicity, we will refer to `log(CA19-9)` as biomarker X_1 and `log(CA125)` as biomarker X_2, and the correlation between these two biomarkers is 0.252. The aim of this study is to identify the best linear combination of X_1 and X_2 that results in a lower OVL value or a higher AUC value compared to using either biomarker alone. Similar to the simulation study, we will be reporting the values of $V = 1 - \text{OVL}$ and $G = 2\text{AUC} - 1$ for ease of comparison.

Biomarkers	$\widehat{\alpha}_b$	$\widehat{G}_b$	$\widehat{\alpha}_k$	$\widehat{G}_k$	$\widehat{\alpha}_e$	$\widehat{G}_e$
X_1		0.764		0.704		0.723
X_2		0.366		0.389		0.411
(X_1, X_2)	(0.789, 0.211)	0.788 (+3%)	(0.694, 0.306)	0.769 (+10%)	(0.719, 0.281)	0.788 (+9%)

Biomarkers	$\widehat{\alpha}_b$	$\widehat{V}_b$	$\widehat{\alpha}_k$	$\widehat{V}_k$	$\widehat{\alpha}_e$	$\widehat{V}_e$
X_1	0.704		0.645		0.658	
X_2	0.280		0.305		0.383	
(X_1, X_2)	(0.847, 0.153)	0.724 (+3%)	(0.840, 0.160)	0.667 (+3%)	(0.613, 0.387)	0.710 (+8%)

Table 1: Pancreatic cancer data set results (Example 1)

Table 1 shows the optimal values of α_1 and α_2 that maximise the values of V and G (minimise the OVL value or maximise the AUC value) using three different approaches: parametric, nonparametric, and

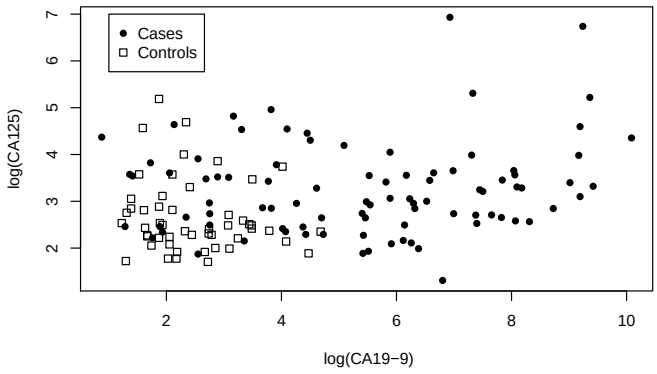

Figure 2: Pancreatic cancer data set (Example 1)

empirical. These approaches are denoted by $(\widehat{G}_b, \widehat{V}_b)$, $(\widehat{G}_k, \widehat{V}_k)$, and $(\widehat{G}_e, \widehat{V}_e)$, respectively. From this table, biomarker X_1 shows better discrimination power in comparison to biomarker X_2. However, combining these biomarkers by giving more weight to biomarker X_1 than biomarker X_2 improves the overall discrimination power. In particular, the combination of biomarker X_1 with a weight of about 70-72% using kernel-based or empirical estimates for G results in the highest improvement of 9% to 10% when compared to using biomarker X_1 alone. On the other hand, the highest improvement of 8% is obtained by combining biomarker X_1 with a weight of about 61% using the empirical estimate for V.

Example 2 (**DMD data set).** In this example, we use a dataset that was first discussed by Cox *et al.* [14]. The dataset was created to develop screening methods to identify carriers of a rare genetic disorder. The dataset comprises four measurements, M_1, M_2, M_3, and M_4, made on blood samples. We use a subset of this dataset which consists of 120 observations, 82 of which are *normals* and 38 are *carriers*. The four measurements were transformed to a natural logarithmic scale and standardized; these are displayed in Figure 3. From the correlation matrix given below, we can observe that M_1 is highly correlated with M_3 and M_4. Similarly, M_3 is also highly correlated with M_4. However, M_2 is only weakly correlated with the other measurements.

$$\text{Corr} = \begin{array}{c} \\ M_1 \\ M_2 \\ M_3 \\ M_4 \end{array} \begin{array}{c} \begin{array}{cccc} M_1 & M_2 & M_3 & M_4 \end{array} \\ \begin{pmatrix} 1.000 & 0.115 & 0.644 & 0.642 \\ 0.115 & 1.000 & 0.221 & 0.284 \\ 0.644 & 0.221 & 1.000 & 0.561 \\ 0.642 & 0.284 & 0.561 & 1.000 \end{pmatrix} \end{array}.$$

As we can see from Tables 2 and 3, measurement M_1 has the highest discrimination power (largest G and V values), followed by M_4, then M_3, while M_2 shows the worst discrimination power across considering all

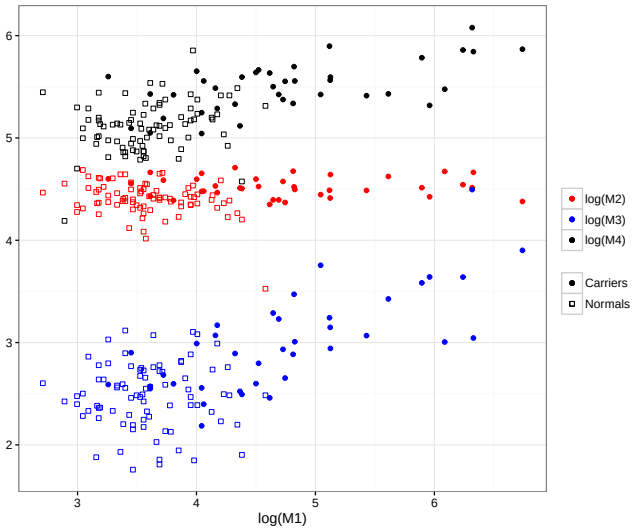

Figure 3: DMD data set (Example 2)

the different estimate values of G and V. Tables 2 and 3 also show the optimal values of α that maximise the values of V and G (minimise the OVL value or maximise the AUC value) using three different approaches: parametric, nonparametric, and empirical. We consider all possible combinations of measurements, that is, combining all possible ways of combining two, three and four measurements.

It's interesting to note that when combining two measurements, the best and worst measurements, M_1 and M_2, (with a correlation coefficient of 0.115), lead to the highest improvement in accuracy, regardless of whether we aim to maximise G or V and regardless of which estimations methods are used. The second best combination is M_1 and M_4, according to estimates of G values. However, there is no agreement between the different estimates of V on which combination should take the second position. V_k agrees with the decision made by G values, but V_b and V_e did not place M_1 and M_4 as their second candidate. Nevertheless, the corresponding values are not too far off. There is an agreement among all methods that combining M_2 and M_3 leads to worse accuracy compared to other combinations. Although the correlation between M_2 and M_3 is not very strong (0.221), this may be because we are combining initially weak measurements, and combining them still provides more discrimination power than considering one of them alone.

Similarly, when considering all possible combinations of three measurements, all methods suggest that combining M_1, M_2, and M_4 will give the best performance, closely followed by combining M_1, M_2, and M_3. Finally, when combining all four measurements, it seems that G_b and G_k gave more weight to M_1, followed by M_2 and M_4, almost neglecting M_3. On the other hand, G_e gave more weight to M_1 and M_2 and almost neglected M_3 and M_4. V_b gave more weight to M_1 and M_2, and smaller weight to M_3 and M_4. V_k and V_e gave more weight to M_1, M_2, and M_4, and almost neglected M_3. As we can see, the improvement gained

by combining all measurements is at least 14% for G values and 12% for V values.

Measurements	$\widehat{\alpha}_b$	$\widehat{G}_b$	$\widehat{\alpha}_k$	$\widehat{G}_k$	$\widehat{\alpha}_e$	$\widehat{G}_e$
M_1		0.779		0.764		0.807
M_2		0.478		0.484		0.505
M_3		0.644		0.606		0.646
M_4		0.752		0.737		0.758
(M_1, M_2)	(0.62, 0.38)	0.861 [11%]	(0.58, 0.42)	0.887 [16%]	(0.61, 0.39)	0.906 [12%]
(M_1, M_3)	(0.76, 0.24)	0.792	(0.70, 0.30)	0.780	(0.75, 0.25)	0.836
(M_1, M_4)	(0.52, 0.48)	0.837	(0.53, 0.47)	0.830	(0.55, 0.45)	0.863
(M_2, M_3)	(0.42, 0.58)	0.741	(0.47, 0.53)	0.737	(0.50, 0.50)	0.756
(M_2, M_4)	(0.32, 0.68)	0.804	(0.35, 0.65)	0.800	(0.31, 0.69)	0.818
(M_3, M_4)	(0.34, 0.66)	0.793	(0.35, 0.65)	0.803	(0.27, 0.73)	0.831
(M_1, M_2, M_3)	(0.53, 0.36, 0.11)	0.865	(0.53, 0.38, 0.09)	0.891	(0.52, 0.34, 0.14)	0.918
(M_1, M_2, M_4)	(0.42, 0.30, 0.28)	0.886 [14%]	(0.47, 0.34, 0.19)	0.902 [18%]	(0.54, 0.32, 0.14)	0.918 [14%]
(M_1, M_3, M_4)	(0.45, 0.11, 0.44)	0.840	(0.50, 0.14, 0.36)	0.839	(0.55, 0.09, 0.36)	0.867
(M_2, M_3, M_4)	(0.28, 0.26, 0.46)	0.837	(0.30, 0.27, 0.43)	0.840	(0.27, 0.28, 0.45)	0.854
(M_1, M_2, M_3, M_4)	(0.39, 0.30, 0.05, 0.260)	0.887 [14%]	(0.42, 0.32, 0.08, 0.180)	0.905 [18%]	(0.52, 0.31, 0.08, 0.09)	0.924 [15%]

Table 2: DMD data set, measurements are combined, $G = 2AUC - 1$

Measurements	$\widehat{\alpha}_b$	$\widehat{V}_b$	$\widehat{\alpha}_k$	$\widehat{V}_k$	$\widehat{\alpha}_e$	$\widehat{V}_e$
$M1$		0.690		0.615		0.696
$M2$		0.408		0.359		0.420
$M3$		0.513		0.443		0.534
$M4$		0.586		0.552		0.633
(M_1, M_2)	(0.74, 0.26)	0.753 [10%]	(0.68, 0.32)	0.706 [15%]	(0.64, 0.36)	0.846[22%]
(M_1, M_3)	(0.74, 0.26)	0.719	(0.74, 0.26)	0.644	(0.84, 0.16)	0.725
(M_1, M_4)	(0.71, 0.29)	0.718	(0.61, 0.39)	0.654	(0.54, 0.46)	0.748
(M_2, M_3)	(0.42, 0.58)	0.575	(0.46, 0.54)	0.537	(0.52, 0.48)	0.704
(M_2, M_4)	(0.36, 0.64)	0.646	(0.34, 0.66)	0.604	(0.62, 0.38)	0.730
(M_3, M_4)	(0.39, 0.61)	0.635	(0.34, 0.66)	0.600	(0.35, 0.65)	0.757
(M_1, M_2, M_3)	(0.64, 0.22, 0.14)	0.765	(0.65, 0.27, 0.08)	0.710	(0.64, 0.36, 0.00)	0.846
(M_1, M_2, M_4)	(0.60, 0.23, 0.17)	0.765 [11%]	(0.48, 0.30, 0.22)	0.722 [17%]	(0.48, 0.40, 0.12)	0.860 [24%]
(M_1, M_3, M_4)	(0.61, 0.19, 0.20)	0.732	(0.54, 0.15, 0.31)	0.658	(0.09, 0.26, 0.65)	0.781
(M_2, M_3, M_4)	(0.29, 0.25, 0.46)	0.677	(0.28, 0.25, 0.47)	0.655	(0.32, 0.28, 0.40)	0.787
(M_1, M_2, M_3, M_4)	(0.56, 0.21, 0.11, 0.12)	0.771 [12%]	(0.44, 0.30, 0.06, 0.20)	0.724 [18%]	(0.48, 0.40, 0.00, 0.12)	0.860 [24%]

Table 3: DMD data set, measurements are combined, $V = 1 - OVL$

6 Concluding remarks

In this paper, we explored how combining two or more biomarkers can improve diagnostic accuracy by reducing the overlap coefficient. We introduced the overlap coefficient as an objective function to minimise and compared its results with those based on the more commonly used AUC measure. Both parametric and non-parametric approaches were employed to estimate the density functions, and we provided a probabilistic interpretation of the overlap measure, offering additional insight into its application. The simulation study demonstrated that combining biomarkers leads to an overall improvement in diagnostic accuracy compared to using a single biomarker. This holds true regardless of whether parametric or non-parametric estimation methods are used.

Our results show that when the individual accuracies of biomarkers are equal, they are assigned equal weights in the combined score. However, when one biomarker has a higher diagnostic accuracy, it is given

more weight. In cases where biomarkers are highly correlated, the weights tend to be skewed towards the biomarker with the higher accuracy. The most significant improvements in diagnostic accuracy are generally observed when two independent biomarkers of equal accuracy are combined. These findings highlight the potential of combining biomarkers for enhanced diagnostic performance. The simulation results presented in this study could also serve as the foundation for a more formal decision-making framework for combining biomarkers, which could help guide the selection of optimal combinations in practice. Future research could further explore this framework, considering additional factors such as the cost of diagnostic tests and practical constraints in clinical settings.

The proposed method for determining the optimal set of weights for combining biomarkers involves evaluating the overlap coefficient (OVL) for various weight combinations. This process can become computationally intensive, especially as the number of biomarkers increases. To mitigate this, future work could focus on developing more efficient optimisation techniques to reduce the computational burden, particularly when working with large datasets or a high number of biomarkers. Exploring alternative search methods or approximations could enhance scalability while maintaining accuracy.

While the overlap measure is effective for comparing biomarkers in the case of normal distributions, its application to non-normal, skewed, or multimodal distributions could present challenges. Future research could explore alternative methods, such as the use of weighted AUC, which offers a more robust evaluation of biomarker accuracy, particularly in complex data structures like those encountered in genetic data [15]. Additionally, the Weitzman overlapping coefficient, as described by Montoya et al. [16], could offer a valuable tool for comparing distributions with varying shapes, providing further insights into biomarker comparison in non-normal data.

Finally, an important avenue for future work is benchmarking the proposed methodology against logistic regression, a widely used method for evaluating biomarker combinations. Such a comparison would offer valuable insights into the relative strengths of our approach, further enhancing its practical relevance in biomarker research.

Acknowledgement

We would like to thank the anonymous reviewers for their helpful comments and constructive feedback, which have improved the quality and clarity of this manuscript.

Conflict of Interest

The author has declared no conflict of interest.

Appendix: Simulation study results

G_1	G_2	$\widehat{\alpha}_b$		$\widehat{G}_b$	$\widehat{\alpha}_k$		$\widehat{G}_k$	$\widehat{\alpha}_e$		$\widehat{G}_e$	$\widehat{\alpha}_b$		$\widehat{V}_b$	$\widehat{\alpha}_k$		$\widehat{V}_k$	$\widehat{\alpha}_e$		$\widehat{V}_e$
$n_0 = n_1 = 50$, $\rho = 0$, $\sigma_1 = 1$, $\sigma_2 = 1$																			
0.2	0.2	0.52	0.48	0.286	0.44	0.56	0.259	0.40	0.60	0.274	0.45	0.55	0.215	0.69	0.31	0.225	0.74	0.26	0.320
0.4	0.2	0.70	0.30	0.453	0.80	0.20	0.424	0.84	0.16	0.439	0.69	0.31	0.330	0.69	0.31	0.310	0.98	0.02	0.400
0.4	0.4	0.53	0.47	0.548	0.54	0.46	0.511	0.53	0.47	0.535	0.51	0.49	0.410	0.60	0.40	0.359	0.59	0.41	0.440
0.6	0.2	0.80	0.20	0.641	0.81	0.19	0.625	0.83	0.17	0.667	0.81	0.19	0.484	0.93	0.07	0.446	0.87	0.13	0.560
0.6	0.4	0.65	0.35	0.692	0.70	0.30	0.670	0.79	0.21	0.703	0.64	0.36	0.530	0.68	0.32	0.472	0.86	0.14	0.580
0.6	0.6	0.54	0.46	0.771	0.59	0.41	0.746	0.65	0.35	0.776	0.53	0.47	0.608	0.58	0.42	0.540	0.66	0.34	0.640
0.8	0.2	0.86	0.14	0.830	0.80	0.20	0.818	0.75	0.25	0.850	0.88	0.12	0.669	0.89	0.11	0.635	0.77	0.23	0.780
0.8	0.4	0.75	0.25	0.851	0.73	0.27	0.843	0.76	0.24	0.867	0.75	0.25	0.692	0.80	0.20	0.653	0.78	0.22	0.800
0.8	0.6	0.65	0.35	0.885	0.66	0.34	0.879	0.66	0.34	0.897	0.64	0.36	0.735	0.70	0.30	0.690	0.80	0.20	0.800
0.8	0.8	0.54	0.46	0.933	0.56	0.44	0.928	0.53	0.47	0.944	0.53	0.47	0.806	0.61	0.39	0.756	0.75	0.25	0.860
$n_0 = n_1 = 50$, $\rho = 0.5$, $\sigma_1 = 1$, $\sigma_2 = 1$																			
0.2	0.2	0.44	0.56	0.239	0.42	0.58	0.239	0.46	0.54	0.258	0.66	0.34	0.179	0.30	0.70	0.206	0.33	0.67	0.300
0.4	0.2	1.00	0.00	0.402	0.82	0.18	0.365	0.98	0.02	0.392	1.00	0.00	0.301	0.97	0.03	0.268	0.99	0.01	0.340
0.4	0.4	0.46	0.54	0.456	0.41	0.59	0.439	0.37	0.63	0.463	0.52	0.48	0.336	0.40	0.60	0.329	0.44	0.56	0.420
0.6	0.2	1.00	0.00	0.598	1.00	0.00	0.547	0.98	0.02	0.577	1.00	0.00	0.454	0.89	0.11	0.388	0.93	0.07	0.460
0.6	0.4	0.88	0.12	0.600	0.78	0.22	0.560	0.71	0.29	0.585	0.89	0.11	0.456	0.73	0.27	0.404	0.71	0.29	0.500
0.6	0.6	0.47	0.53	0.663	0.41	0.59	0.636	0.36	0.64	0.661	0.50	0.50	0.506	0.42	0.58	0.471	0.59	0.41	0.560
0.8	0.2	1.00	0.00	0.796	1.00	0.00	0.748	0.99	0.01	0.784	1.00	0.00	0.636	1.00	0.00	0.562	1.00	0.00	0.660
0.8	0.4	1.00	0.00	0.796	1.00	0.00	0.748	0.99	0.01	0.784	1.00	0.00	0.636	1.00	0.00	0.562	0.86	0.14	0.700
0.8	0.6	0.84	0.16	0.800	0.78	0.22	0.762	0.87	0.13	0.800	0.85	0.15	0.640	0.70	0.30	0.582	0.86	0.14	0.720
0.8	0.8	0.47	0.53	0.854	0.42	0.58	0.823	0.34	0.66	0.853	0.49	0.51	0.697	0.44	0.56	0.659	0.28	0.72	0.800
$n_0 = n_1 = 50$, $\rho = 0.75$, $\sigma_1 = 1$, $\sigma_2 = 1$																			
0.2	0.2	0.41	0.59	0.223	0.39	0.61	0.223	0.43	0.57	0.245	0.79	0.21	0.169	0.20	0.80	0.198	0.38	0.62	0.280
0.4	0.2	1.00	0.00	0.400	1.00	0.00	0.370	1.00	0.00	0.384	1.00	0.00	0.300	1.00	0.00	0.267	0.51	0.49	0.320
0.4	0.4	0.44	0.56	0.426	0.36	0.64	0.411	0.31	0.69	0.434	0.55	0.45	0.314	0.34	0.66	0.313	0.38	0.62	0.400
0.6	0.2	1.00	0.00	0.595	1.00	0.00	0.552	1.00	0.00	0.573	1.00	0.00	0.451	1.00	0.00	0.399	1.00	0.00	0.480
0.6	0.4	1.00	0.00	0.595	1.00	0.00	0.552	0.92	0.08	0.577	1.00	0.00	0.451	1.00	0.00	0.399	1.00	0.00	0.480
0.6	0.6	0.45	0.55	0.627	0.36	0.64	0.600	0.42	0.58	0.629	0.50	0.50	0.474	0.37	0.63	0.442	0.44	0.56	0.520
0.8	0.2	1.00	0.00	0.792	1.00	0.00	0.750	1.00	0.00	0.786	1.00	0.00	0.632	1.00	0.00	0.570	1.00	0.00	0.700
0.8	0.4	1.00	0.00	0.792	1.00	0.00	0.750	1.00	0.00	0.786	1.00	0.00	0.632	1.00	0.00	0.570	1.00	0.00	0.700
0.8	0.6	1.00	0.00	0.792	1.00	0.00	0.750	1.00	0.00	0.786	1.00	0.00	0.632	1.00	0.00	0.570	1.00	0.00	0.700
0.8	0.8	0.46	0.54	0.822	0.38	0.62	0.790	0.41	0.59	0.827	0.49	0.51	0.661	0.40	0.60	0.622	0.45	0.55	0.760

Table 4: Simulated data example, Bivariate normal distribution, Scenario 1, $n_0 = n_1 = 50$

G_1	G_2	$\widehat{\alpha}_b$		$\widehat{G}_b$	$\widehat{\alpha}_k$		$\widehat{G}_k$	$\widehat{\alpha}_e$		$\widehat{G}_e$	$\widehat{\alpha}_b$		$\widehat{V}_b$	$\widehat{\alpha}_k$		$\widehat{V}_k$	$\widehat{\alpha}_e$		$\widehat{V}_e$
$n_0 = n_1 = 100$, $\rho = 0$, $\sigma_1 = 1$, $\sigma_2 = 1$																			
0.2	0.2	0.50	0.50	0.291	0.55	0.45	0.318	0.52	0.48	0.330	0.47	0.53	0.221	0.54	0.46	0.239	0.38	0.62	0.320
0.4	0.2	0.65	0.35	0.443	0.59	0.41	0.466	0.60	0.40	0.485	0.62	0.38	0.327	0.62	0.38	0.345	0.69	0.31	0.410
0.4	0.4	0.50	0.50	0.545	0.52	0.48	0.566	0.53	0.47	0.587	0.49	0.51	0.410	0.51	0.49	0.425	0.48	0.52	0.530
0.6	0.2	0.73	0.27	0.614	0.67	0.33	0.622	0.71	0.29	0.643	0.71	0.29	0.462	0.66	0.34	0.467	0.48	0.52	0.540
0.6	0.4	0.60	0.40	0.677	0.57	0.43	0.689	0.57	0.43	0.709	0.59	0.41	0.519	0.56	0.44	0.531	0.52	0.48	0.630
0.6	0.6	0.50	0.50	0.764	0.53	0.47	0.765	0.53	0.47	0.782	0.50	0.50	0.603	0.52	0.48	0.612	0.55	0.45	0.690
0.8	0.2	0.78	0.22	0.796	0.72	0.28	0.786	0.71	0.29	0.807	0.78	0.22	0.632	0.67	0.33	0.620	0.69	0.31	0.710
0.8	0.4	0.68	0.32	0.827	0.65	0.35	0.819	0.65	0.35	0.838	0.67	0.33	0.666	0.60	0.40	0.665	0.70	0.30	0.740
0.8	0.6	0.59	0.41	0.871	0.58	0.42	0.860	0.58	0.42	0.870	0.58	0.42	0.719	0.56	0.44	0.719	0.58	0.42	0.800
0.8	0.8	0.50	0.50	0.927	0.47	0.53	0.912	0.51	0.49	0.920	0.50	0.50	0.798	0.52	0.48	0.786	0.62	0.38	0.850
$n_0 = n_1 = 100$, $\rho = 0.5$, $\sigma_1 = 1$, $\sigma_2 = 1$																			
0.2	0.2	0.62	0.38	0.229	0.66	0.34	0.235	0.73	0.27	0.252	0.75	0.25	0.177	0.73	0.27	0.189	1.00	0.00	0.250
0.4	0.2	1.00	0.00	0.404	1.00	0.00	0.414	0.97	0.03	0.429	1.00	0.00	0.303	1.00	0.00	0.313	0.99	0.01	0.380
0.4	0.4	0.59	0.41	0.441	0.63	0.37	0.433	0.68	0.32	0.456	0.63	0.37	0.327	0.70	0.30	0.326	0.71	0.29	0.390
0.6	0.2	1.00	0.00	0.598	1.00	0.00	0.600	1.00	0.00	0.624	1.00	0.00	0.454	1.00	0.00	0.455	1.00	0.00	0.540
0.6	0.4	0.92	0.08	0.599	1.00	0.00	0.600	1.00	0.00	0.624	0.93	0.07	0.455	1.00	0.00	0.455	1.00	0.00	0.540
0.6	0.6	0.58	0.42	0.647	0.62	0.38	0.627	0.56	0.44	0.652	0.60	0.40	0.493	0.68	0.32	0.471	0.98	0.02	0.550
0.8	0.2	1.00	0.00	0.794	1.00	0.00	0.785	1.00	0.00	0.802	1.00	0.00	0.634	1.00	0.00	0.617	1.00	0.00	0.690
0.8	0.4	1.00	0.00	0.794	1.00	0.00	0.785	1.00	0.00	0.802	1.00	0.00	0.634	1.00	0.00	0.617	1.00	0.00	0.690
0.8	0.6	0.88	0.12	0.797	0.94	0.06	0.785	0.90	0.10	0.805	0.89	0.11	0.637	1.00	0.00	0.617	1.00	0.00	0.690
0.8	0.8	0.58	0.42	0.840	0.62	0.38	0.813	0.66	0.34	0.838	0.59	0.41	0.682	0.65	0.35	0.643	0.33	0.67	0.710
$n_0 = n_1 = 100$, $\rho = 0.75$, $\sigma_1 = 1$, $\sigma_2 = 1$																			
0.2	0.2	0.68	0.32	0.213	0.90	0.10	0.219	0.84	0.16	0.236	0.92	0.08	0.167	0.86	0.14	0.179	0.55	0.45	0.230
0.4	0.2	1.00	0.00	0.400	1.00	0.00	0.400	1.00	0.00	0.419	1.00	0.00	0.299	1.00	0.00	0.305	1.00	0.00	0.360
0.4	0.4	0.64	0.36	0.412	0.75	0.25	0.404	0.86	0.14	0.428	0.71	0.29	0.306	0.81	0.19	0.307	0.91	0.09	0.370
0.6	0.2	1.00	0.00	0.593	1.00	0.00	0.583	1.00	0.00	0.608	1.00	0.00	0.449	1.00	0.00	0.441	1.00	0.00	0.520
0.6	0.4	1.00	0.00	0.593	1.00	0.00	0.583	1.00	0.00	0.608	1.00	0.00	0.449	1.00	0.00	0.441	1.00	0.00	0.520
0.6	0.6	0.63	0.37	0.610	0.62	0.38	0.590	0.67	0.33	0.620	0.66	0.34	0.462	0.78	0.22	0.444	1.00	0.00	0.520
0.8	0.2	1.00	0.00	0.789	1.00	0.00	0.771	1.00	0.00	0.795	1.00	0.00	0.629	1.00	0.00	0.598	1.00	0.00	0.640
0.8	0.4	1.00	0.00	0.789	1.00	0.00	0.771	1.00	0.00	0.795	1.00	0.00	0.629	1.00	0.00	0.598	1.00	0.00	0.640
0.8	0.6	1.00	0.00	0.789	1.00	0.00	0.771	1.00	0.00	0.795	1.00	0.00	0.629	1.00	0.00	0.598	1.00	0.00	0.640
0.8	0.8	0.62	0.38	0.807	0.62	0.38	0.778	0.81	0.19	0.804	0.64	0.36	0.646	0.69	0.31	0.609	0.31	0.69	0.680

Table 5: Simulated data example, Bivariate normal distribution, Scenario 1, $n_0 = n_1 = 100$

G_1	G_2	$\widehat{\alpha}_b$		$\widehat{G}_b$	$\widehat{\alpha}_k$		$\widehat{G}_k$	$\widehat{\alpha}_e$		$\widehat{G}_e$	$\widehat{\alpha}_b$		$\widehat{V}_b$	$\widehat{\alpha}_k$		$\widehat{V}_k$	$\widehat{\alpha}_e$		$\widehat{V}_e$
$n_0 = n_1 = 50$, $\rho = 0$, $\sigma_1 = 2$, $\sigma_2 = 1$																			
0.2	0.2	0.39	0.61	0.286	0.33	0.67	0.268	0.29	0.71	0.290	1.00	0.00	0.333	0.95	0.05	0.369	0.46	0.54	0.360
0.4	0.2	0.56	0.44	0.444	0.56	0.44	0.402	0.75	0.25	0.440	0.90	0.10	0.405	0.84	0.16	0.428	0.54	0.46	0.440
0.4	0.4	0.40	0.60	0.545	0.34	0.66	0.514	0.39	0.61	0.548	0.63	0.37	0.422	0.82	0.18	0.436	0.58	0.42	0.520
0.6	0.2	0.66	0.34	0.623	0.63	0.37	0.571	0.69	0.31	0.620	0.83	0.17	0.524	0.89	0.11	0.513	0.96	0.04	0.580
0.6	0.4	0.51	0.49	0.682	0.43	0.57	0.634	0.52	0.48	0.662	0.64	0.36	0.549	0.77	0.23	0.523	0.63	0.37	0.660
0.6	0.6	0.40	0.60	0.767	0.35	0.65	0.733	0.33	0.67	0.758	0.46	0.54	0.610	0.49	0.51	0.573	0.49	0.51	0.700
0.8	0.2	0.74	0.26	0.809	0.83	0.17	0.755	0.76	0.24	0.791	0.84	0.16	0.687	0.80	0.20	0.647	0.81	0.19	0.760
0.8	0.4	0.60	0.40	0.835	0.56	0.44	0.790	0.54	0.46	0.817	0.70	0.30	0.705	0.69	0.31	0.665	0.84	0.16	0.760
0.8	0.6	0.50	0.50	0.876	0.44	0.56	0.845	0.46	0.54	0.884	0.57	0.43	0.740	0.59	0.41	0.697	0.68	0.32	0.780
0.8	0.8	0.40	0.60	0.930	0.41	0.59	0.914	0.41	0.59	0.939	0.43	0.57	0.804	0.45	0.55	0.756	0.47	0.53	0.860
$n_0 = n_1 = 50$, $\rho = 0.5$, $\sigma_1 = 2$, $\sigma_2 = 1$																			
0.2	0.2	0.39	0.61	0.262	0.10	0.90	0.245	0.58	0.42	0.272	1.00	0.00	0.325	1.00	0.00	0.361	1.00	0.00	0.380
0.4	0.2	0.90	0.10	0.421	0.85	0.15	0.395	0.76	0.24	0.428	1.00	0.00	0.402	1.00	0.00	0.418	1.00	0.00	0.440
0.4	0.4	0.38	0.62	0.483	0.35	0.65	0.443	0.46	0.54	0.482	1.00	0.00	0.402	1.00	0.00	0.418	0.26	0.74	0.460
0.6	0.2	1.00	0.00	0.612	0.95	0.05	0.565	0.98	0.02	0.600	1.00	0.00	0.521	1.00	0.00	0.517	1.00	0.00	0.600
0.6	0.4	0.71	0.29	0.622	0.75	0.25	0.577	0.77	0.23	0.618	1.00	0.00	0.521	1.00	0.00	0.517	1.00	0.00	0.600
0.6	0.6	0.38	0.62	0.689	0.35	0.65	0.637	0.39	0.61	0.670	0.55	0.45	0.540	0.91	0.09	0.518	0.94	0.06	0.620
0.8	0.2	1.00	0.00	0.805	1.00	0.00	0.752	1.00	0.00	0.805	1.00	0.00	0.684	1.00	0.00	0.653	1.00	0.00	0.740
0.8	0.4	1.00	0.00	0.805	0.93	0.07	0.753	1.00	0.00	0.805	1.00	0.00	0.684	1.00	0.00	0.653	1.00	0.00	0.740
0.8	0.6	0.67	0.33	0.817	0.74	0.26	0.766	0.93	0.07	0.806	0.88	0.12	0.685	1.00	0.00	0.653	1.00	0.00	0.740
0.8	0.8	0.38	0.62	0.872	0.34	0.66	0.828	0.32	0.68	0.860	0.46	0.54	0.726	0.67	0.33	0.665	0.69	0.31	0.760
$n_0 = n_1 = 50$, $\rho = 0.75$, $\sigma_1 = 2$, $\sigma_2 = 1$																			
0.2	0.2	0.38	0.62	0.249	0.56	0.44	0.239	0.54	0.46	0.266	1.00	0.00	0.325	1.00	0.00	0.351	1.00	0.00	0.360
0.4	0.2	1.00	0.00	0.423	1.00	0.00	0.395	1.00	0.00	0.428	1.00	0.00	0.403	1.00	0.00	0.414	1.00	0.00	0.440
0.4	0.4	0.37	0.63	0.455	0.56	0.44	0.420	0.38	0.62	0.464	1.00	0.00	0.403	1.00	0.00	0.414	1.00	0.00	0.440
0.6	0.2	1.00	0.00	0.615	1.00	0.00	0.568	1.00	0.00	0.602	1.00	0.00	0.523	1.00	0.00	0.510	1.00	0.00	0.580
0.6	0.4	1.00	0.00	0.615	0.95	0.05	0.568	1.00	0.00	0.602	1.00	0.00	0.523	1.00	0.00	0.510	1.00	0.00	0.580
0.6	0.6	0.37	0.63	0.655	0.24	0.76	0.602	0.51	0.49	0.643	1.00	0.00	0.523	1.00	0.00	0.510	1.00	0.00	0.580
0.8	0.2	1.00	0.00	0.806	1.00	0.00	0.753	1.00	0.00	0.799	1.00	0.00	0.685	1.00	0.00	0.650	1.00	0.00	0.720
0.8	0.4	1.00	0.00	0.806	1.00	0.00	0.753	1.00	0.00	0.799	1.00	0.00	0.685	1.00	0.00	0.650	1.00	0.00	0.720
0.8	0.6	1.00	0.00	0.806	0.94	0.06	0.753	1.00	0.00	0.799	1.00	0.00	0.685	1.00	0.00	0.650	1.00	0.00	0.720
0.8	0.8	0.37	0.63	0.842	0.24	0.76	0.791	0.39	0.61	0.829	0.60	0.40	0.693	1.00	0.00	0.650	0.95	0.05	0.740

Table 6: Simulated data example, Bivariate normal distribution, Scenario 2, $n_0 = n_1 = 50$

G_1	G_2	$\widehat{\alpha}_b$		$\widehat{G}_b$	$\widehat{\alpha}_k$		$\widehat{G}_k$	$\widehat{\alpha}_e$		$\widehat{G}_e$	$\widehat{\alpha}_b$		$\widehat{V}_b$	$\widehat{\alpha}_k$		$\widehat{V}_k$	$\widehat{\alpha}_e$		$\widehat{V}_e$
$n_0 = n_1 = 100$, $\rho = 0$, $\sigma_1 = 2$, $\sigma_2 = 1$																			
0.2	0.2	0.38	0.62	0.278	0.55	0.45	0.297	0.62	0.38	0.318	0.89	0.11	0.314	0.69	0.31	0.329	0.64	0.36	0.350
0.4	0.2	0.58	0.42	0.425	0.68	0.32	0.445	0.65	0.35	0.466	0.84	0.16	0.388	0.72	0.28	0.414	0.72	0.28	0.450
0.4	0.4	0.38	0.62	0.519	0.44	0.56	0.511	0.41	0.59	0.541	0.60	0.40	0.411	0.64	0.36	0.439	0.67	0.33	0.490
0.6	0.2	0.71	0.29	0.597	0.78	0.22	0.595	0.71	0.29	0.616	0.83	0.17	0.502	0.99	0.01	0.519	0.70	0.30	0.560
0.6	0.4	0.51	0.49	0.649	0.55	0.45	0.635	0.53	0.47	0.659	0.64	0.36	0.526	0.70	0.30	0.527	0.61	0.39	0.590
0.6	0.6	0.38	0.62	0.735	0.37	0.63	0.708	0.35	0.65	0.731	0.45	0.55	0.584	0.56	0.44	0.564	0.45	0.55	0.630
0.8	0.2	0.82	0.18	0.784	0.84	0.16	0.753	0.84	0.16	0.773	0.87	0.13	0.660	1.00	0.00	0.657	1.00	0.00	0.730
0.8	0.4	0.63	0.37	0.805	0.67	0.33	0.771	0.60	0.40	0.788	0.71	0.29	0.675	0.85	0.15	0.662	1.00	0.00	0.730
0.8	0.6	0.50	0.50	0.846	0.49	0.51	0.810	0.45	0.55	0.826	0.57	0.43	0.709	0.75	0.25	0.679	0.62	0.38	0.750
0.8	0.8	0.38	0.62	0.907	0.35	0.65	0.880	0.30	0.70	0.897	0.42	0.58	0.774	0.43	0.57	0.727	0.44	0.56	0.820
$n_0 = n_1 = 100$, $\rho = 0.5$, $\sigma_1 = 2$, $\sigma_2 = 1$																			
0.2	0.2	0.45	0.55	0.280	0.64	0.36	0.308	0.73	0.27	0.332	1.00	0.00	0.332	1.00	0.00	0.344	1.00	0.00	0.350
0.4	0.2	0.97	0.03	0.438	0.93	0.07	0.457	0.87	0.13	0.481	1.00	0.00	0.411	1.00	0.00	0.429	0.74	0.26	0.450
0.4	0.4	0.43	0.57	0.487	0.49	0.51	0.495	0.49	0.51	0.522	0.87	0.13	0.413	0.81	0.19	0.431	0.49	0.51	0.480
0.6	0.2	1.00	0.00	0.622	1.00	0.00	0.615	1.00	0.00	0.632	1.00	0.00	0.529	1.00	0.00	0.534	1.00	0.00	0.600
0.6	0.4	0.78	0.22	0.627	0.72	0.28	0.623	0.64	0.36	0.651	0.98	0.02	0.529	0.86	0.14	0.536	1.00	0.00	0.600
0.6	0.6	0.42	0.58	0.683	0.48	0.52	0.670	0.51	0.49	0.691	0.60	0.40	0.548	0.63	0.37	0.562	0.66	0.34	0.650
0.8	0.2	1.00	0.00	0.807	1.00	0.00	0.774	1.00	0.00	0.795	1.00	0.00	0.686	1.00	0.00	0.673	1.00	0.00	0.770
0.8	0.4	1.00	0.00	0.807	1.00	0.00	0.774	0.99	0.01	0.795	1.00	0.00	0.686	0.88	0.12	0.675	1.00	0.00	0.770
0.8	0.6	0.74	0.26	0.814	0.64	0.36	0.782	0.63	0.37	0.797	0.87	0.13	0.688	0.79	0.21	0.684	0.91	0.09	0.780
0.8	0.8	0.42	0.58	0.861	0.34	0.66	0.829	0.37	0.63	0.847	0.52	0.48	0.722	0.56	0.44	0.715	0.56	0.44	0.800
$n_0 = n_1 = 100$, $\rho = 0.75$, $\sigma_1 = 2$, $\sigma_2 = 1$																			
0.2	0.2	0.45	0.55	0.272	0.65	0.35	0.299	0.84	0.16	0.323	1.00	0.00	0.340	1.00	0.00	0.348	0.81	0.19	0.350
0.4	0.2	1.00	0.00	0.443	1.00	0.00	0.461	0.97	0.03	0.485	1.00	0.00	0.418	1.00	0.00	0.437	1.00	0.00	0.430
0.4	0.4	0.44	0.56	0.467	0.55	0.45	0.478	0.56	0.44	0.503	1.00	0.00	0.418	1.00	0.00	0.437	0.89	0.11	0.440
0.6	0.2	1.00	0.00	0.627	1.00	0.00	0.620	1.00	0.00	0.637	1.00	0.00	0.535	1.00	0.00	0.546	1.00	0.00	0.610
0.6	0.4	1.00	0.00	0.627	1.00	0.00	0.620	0.89	0.11	0.639	1.00	0.00	0.535	1.00	0.00	0.546	1.00	0.00	0.610
0.6	0.6	0.44	0.56	0.656	0.52	0.48	0.647	0.53	0.47	0.674	0.96	0.04	0.535	0.86	0.14	0.547	0.63	0.37	0.630
0.8	0.2	1.00	0.00	0.811	1.00	0.00	0.777	1.00	0.00	0.795	1.00	0.00	0.691	1.00	0.00	0.686	1.00	0.00	0.770
0.8	0.4	1.00	0.00	0.811	1.00	0.00	0.777	1.00	0.00	0.795	1.00	0.00	0.691	1.00	0.00	0.686	1.00	0.00	0.770
0.8	0.6	1.00	0.00	0.811	1.00	0.00	0.777	1.00	0.00	0.795	1.00	0.00	0.691	1.00	0.00	0.686	1.00	0.00	0.770
0.8	0.8	0.43	0.57	0.837	0.36	0.64	0.806	0.28	0.72	0.824	0.69	0.31	0.698	0.69	0.31	0.693	1.00	0.00	0.770

Table 7: Simulated data example, Bivariate normal distribution, Scenario 2, $n_0 = n_1 = 100$

G_1	G_2	$\widehat{\alpha}_b$		$\widehat{G}_b$	$\widehat{\alpha}_k$		$\widehat{G}_k$	$\widehat{\alpha}_e$		$\widehat{G}_e$	$\widehat{\alpha}_b$		$\widehat{V}_b$	$\widehat{\alpha}_k$		$\widehat{V}_k$	$\widehat{\alpha}_e$		$\widehat{V}_e$
$n_0 = n_1 = 50$, $\rho = 0$, $\sigma_1 = 2$, $\sigma_2 = 1.5$																			
0.2	0.2	0.48	0.52	0.329	0.50	0.50	0.280	0.41	0.59	0.287	1.00	0.00	0.409	1.00	0.00	0.433	0.98	0.02	0.400
0.4	0.2	0.58	0.42	0.497	0.52	0.48	0.444	0.52	0.48	0.476	0.97	0.03	0.467	1.00	0.00	0.465	0.87	0.13	0.500
0.4	0.4	0.49	0.51	0.601	0.47	0.53	0.562	0.44	0.56	0.598	0.56	0.44	0.497	0.83	0.17	0.471	0.64	0.36	0.540
0.6	0.2	0.64	0.36	0.673	0.55	0.45	0.619	0.60	0.40	0.662	0.76	0.24	0.578	0.99	0.01	0.549	0.65	0.35	0.620
0.6	0.4	0.55	0.45	0.739	0.51	0.49	0.707	0.55	0.45	0.742	0.61	0.39	0.618	0.71	0.29	0.552	0.69	0.31	0.640
0.6	0.6	0.49	0.51	0.817	0.46	0.54	0.802	0.45	0.55	0.834	0.52	0.48	0.686	0.49	0.51	0.622	0.58	0.42	0.700
0.8	0.2	0.68	0.32	0.845	0.61	0.39	0.803	0.64	0.36	0.857	0.75	0.25	0.734	0.75	0.25	0.665	0.80	0.20	0.740
0.8	0.4	0.61	0.39	0.877	0.55	0.45	0.856	0.67	0.33	0.895	0.66	0.34	0.763	0.65	0.35	0.696	0.73	0.27	0.780
0.8	0.6	0.55	0.45	0.915	0.51	0.49	0.908	0.53	0.47	0.934	0.58	0.42	0.805	0.55	0.45	0.747	0.48	0.52	0.840
0.8	0.8	0.49	0.51	0.956	0.46	0.54	0.957	0.55	0.45	0.970	0.51	0.49	0.864	0.48	0.52	0.826	0.55	0.45	0.920
$n_0 = n_1 = 50$, $\rho = 0.5$, $\sigma_1 = 2$, $\sigma_2 = 1.5$																			
0.2	0.2	0.15	0.85	0.255	0.17	0.83	0.199	0.16	0.84	0.199	1.00	0.00	0.399	1.00	0.00	0.393	0.73	0.27	0.380
0.4	0.2	0.55	0.45	0.354	0.83	0.17	0.299	0.76	0.24	0.327	1.00	0.00	0.440	1.00	0.00	0.428	1.00	0.00	0.460
0.4	0.4	0.23	0.77	0.483	0.17	0.83	0.422	0.30	0.70	0.460	1.00	0.00	0.440	1.00	0.00	0.428	1.00	0.00	0.460
0.6	0.2	0.85	0.15	0.523	1.00	0.00	0.475	1.00	0.00	0.501	1.00	0.00	0.519	1.00	0.00	0.494	1.00	0.00	0.560
0.6	0.4	0.48	0.52	0.580	0.42	0.58	0.508	0.38	0.62	0.554	1.00	0.00	0.519	0.85	0.15	0.499	1.00	0.00	0.560
0.6	0.6	0.27	0.73	0.698	0.17	0.83	0.646	0.12	0.88	0.686	0.34	0.66	0.559	0.86	0.14	0.510	0.20	0.80	0.600
0.8	0.2	1.00	0.00	0.731	1.00	0.00	0.686	0.98	0.02	0.727	1.00	0.00	0.653	1.00	0.00	0.621	1.00	0.00	0.640
0.8	0.4	0.73	0.27	0.742	0.86	0.14	0.688	0.94	0.06	0.730	1.00	0.00	0.653	1.00	0.00	0.621	1.00	0.00	0.640
0.8	0.6	0.48	0.52	0.792	0.42	0.58	0.738	0.40	0.60	0.786	0.64	0.36	0.666	1.00	0.00	0.621	0.40	0.60	0.680
0.8	0.8	0.28	0.72	0.882	0.17	0.83	0.860	0.23	0.77	0.899	0.33	0.67	0.749	0.27	0.73	0.699	0.38	0.62	0.780
$n_0 = n_1 = 50$, $\rho = 0.75$, $\sigma_1 = 2$, $\sigma_2 = 1.5$																			
0.2	0.2	0.00	1.00	0.228	0.00	1.00	0.175	0.51	0.49	0.177	1.00	0.00	0.387	1.00	0.00	0.389	0.79	0.21	0.360
0.4	0.2	1.00	0.00	0.334	1.00	0.00	0.297	1.00	0.00	0.322	1.00	0.00	0.431	1.00	0.00	0.425	1.00	0.00	0.460
0.4	0.4	0.08	0.92	0.441	0.01	0.99	0.376	0.08	0.92	0.410	1.00	0.00	0.431	1.00	0.00	0.425	1.00	0.00	0.460
0.6	0.2	1.00	0.00	0.530	1.00	0.00	0.478	1.00	0.00	0.493	1.00	0.00	0.515	1.00	0.00	0.494	1.00	0.00	0.560
0.6	0.4	0.71	0.29	0.537	0.97	0.03	0.478	0.90	0.10	0.496	1.00	0.00	0.515	0.92	0.08	0.494	1.00	0.00	0.560
0.6	0.6	0.13	0.87	0.649	0.03	0.97	0.585	0.08	0.92	0.635	0.33	0.67	0.517	0.93	0.07	0.500	0.32	0.68	0.580
0.8	0.2	1.00	0.00	0.739	1.00	0.00	0.690	1.00	0.00	0.730	1.00	0.00	0.654	1.00	0.00	0.613	1.00	0.00	0.640
0.8	0.4	1.00	0.00	0.739	1.00	0.00	0.690	1.00	0.00	0.730	1.00	0.00	0.654	1.00	0.00	0.613	1.00	0.00	0.640
0.8	0.6	0.68	0.32	0.748	0.91	0.09	0.692	0.97	0.03	0.734	1.00	0.00	0.654	1.00	0.00	0.613	1.00	0.00	0.640
0.8	0.8	0.16	0.84	0.844	0.03	0.97	0.803	0.18	0.82	0.850	0.26	0.74	0.703	0.17	0.83	0.644	0.20	0.80	0.760

Table 8: Simulated data example, Bivariate normal distribution, Scenario 3, $n_0 = n_1 = 50$

G_1	G_2	$\widehat{\alpha}_b$		$\widehat{G}_b$	$\widehat{\alpha}_k$		$\widehat{G}_k$	$\widehat{\alpha}_e$		$\widehat{G}_e$	$\widehat{\alpha}_b$		$\widehat{V}_b$	$\widehat{\alpha}_k$		$\widehat{V}_k$	$\widehat{\alpha}_e$		$\widehat{V}_e$
$n_0 = n_1 = 100$, $\rho = 0$, $\sigma_1 = 2$, $\sigma_2 = 1.5$																			
0.2	0.2	0.46	0.54	0.274	0.56	0.44	0.285	0.61	0.39	0.310	0.96	0.04	0.352	0.84	0.16	0.371	0.67	0.33	0.370
0.4	0.2	0.69	0.31	0.435	0.69	0.31	0.441	0.74	0.26	0.467	0.89	0.11	0.424	0.80	0.20	0.450	0.64	0.36	0.490
0.4	0.4	0.47	0.53	0.523	0.51	0.49	0.513	0.49	0.51	0.535	0.62	0.38	0.449	0.70	0.30	0.475	0.82	0.18	0.520
0.6	0.2	0.82	0.18	0.620	0.74	0.26	0.599	0.83	0.17	0.623	0.90	0.10	0.540	0.80	0.20	0.554	0.83	0.17	0.620
0.6	0.4	0.62	0.38	0.664	0.61	0.39	0.640	0.59	0.41	0.658	0.70	0.30	0.561	0.74	0.26	0.576	0.77	0.23	0.640
0.6	0.6	0.47	0.53	0.743	0.49	0.51	0.716	0.42	0.58	0.737	0.53	0.47	0.616	0.61	0.39	0.611	0.69	0.31	0.670
0.8	0.2	0.91	0.09	0.812	0.96	0.04	0.766	1.00	0.00	0.784	0.94	0.06	0.700	0.93	0.07	0.686	1.00	0.00	0.750
0.8	0.4	0.74	0.26	0.827	0.69	0.31	0.787	0.80	0.20	0.802	0.79	0.21	0.712	0.78	0.22	0.696	0.63	0.37	0.780
0.8	0.6	0.60	0.40	0.861	0.55	0.45	0.826	0.55	0.45	0.842	0.65	0.35	0.742	0.66	0.34	0.721	0.61	0.39	0.790
0.8	0.8	0.47	0.53	0.914	0.43	0.57	0.889	0.42	0.58	0.905	0.51	0.49	0.801	0.53	0.47	0.770	0.59	0.41	0.830
$n_0 = n_1 = 100$, $\rho = 0.5$, $\sigma_1 = 2$, $\sigma_2 = 1.5$																			
0.2	0.2	0.40	0.60	0.243	0.40	0.60	0.264	0.46	0.54	0.287	1.00	0.00	0.298	1.00	0.00	0.339	0.26	0.74	0.360
0.4	0.2	0.99	0.01	0.408	0.87	0.13	0.402	1.00	0.00	0.428	1.00	0.00	0.380	1.00	0.00	0.410	0.56	0.44	0.430
0.4	0.4	0.45	0.55	0.464	0.40	0.60	0.460	0.33	0.67	0.483	0.52	0.48	0.401	0.47	0.53	0.428	0.53	0.47	0.500
0.6	0.2	1.00	0.00	0.610	1.00	0.00	0.581	0.99	0.01	0.611	1.00	0.00	0.510	1.00	0.00	0.514	1.00	0.00	0.550
0.6	0.4	0.83	0.17	0.614	0.87	0.13	0.588	0.95	0.05	0.614	0.89	0.11	0.511	1.00	0.00	0.514	0.84	0.16	0.570
0.6	0.6	0.47	0.53	0.674	0.37	0.63	0.642	0.48	0.52	0.664	0.49	0.51	0.553	0.49	0.51	0.546	0.62	0.38	0.620
0.8	0.2	1.00	0.00	0.809	1.00	0.00	0.767	1.00	0.00	0.791	1.00	0.00	0.682	1.00	0.00	0.656	1.00	0.00	0.720
0.8	0.4	1.00	0.00	0.809	1.00	0.00	0.767	1.00	0.00	0.791	1.00	0.00	0.682	1.00	0.00	0.656	1.00	0.00	0.720
0.8	0.6	0.81	0.19	0.814	0.87	0.13	0.774	0.83	0.17	0.795	0.83	0.17	0.686	0.93	0.07	0.656	0.84	0.16	0.730
0.8	0.8	0.48	0.52	0.862	0.43	0.57	0.821	0.53	0.47	0.841	0.49	0.51	0.735	0.53	0.47	0.688	0.76	0.24	0.760
$n_0 = n_1 = 100$, $\rho = 0.75$, $\sigma_1 = 2$, $\sigma_2 = 1.5$																			
0.2	0.2	0.37	0.63	0.227	0.30	0.70	0.248	0.51	0.49	0.273	1.00	0.00	0.303	1.00	0.00	0.339	0.11	0.89	0.340
0.4	0.2	1.00	0.00	0.410	1.00	0.00	0.404	1.00	0.00	0.428	1.00	0.00	0.384	1.00	0.00	0.410	1.00	0.00	0.420
0.4	0.4	0.44	0.56	0.436	0.38	0.62	0.433	0.49	0.51	0.461	0.82	0.18	0.385	1.00	0.00	0.410	0.49	0.51	0.480
0.6	0.2	1.00	0.00	0.611	1.00	0.00	0.584	1.00	0.00	0.610	1.00	0.00	0.512	1.00	0.00	0.512	1.00	0.00	0.540
0.6	0.4	1.00	0.00	0.611	1.00	0.00	0.584	1.00	0.00	0.610	1.00	0.00	0.512	1.00	0.00	0.512	0.93	0.07	0.550
0.6	0.6	0.47	0.53	0.640	0.38	0.62	0.609	0.34	0.66	0.632	0.58	0.42	0.525	0.48	0.52	0.522	0.57	0.43	0.590
0.8	0.2	1.00	0.00	0.809	1.00	0.00	0.768	1.00	0.00	0.788	1.00	0.00	0.683	1.00	0.00	0.652	1.00	0.00	0.700
0.8	0.4	1.00	0.00	0.809	1.00	0.00	0.768	1.00	0.00	0.788	1.00	0.00	0.683	1.00	0.00	0.652	1.00	0.00	0.700
0.8	0.6	1.00	0.00	0.809	1.00	0.00	0.768	1.00	0.00	0.788	1.00	0.00	0.683	1.00	0.00	0.652	0.94	0.06	0.710
0.8	0.8	0.49	0.51	0.834	0.55	0.45	0.791	0.45	0.55	0.811	0.55	0.45	0.702	0.60	0.40	0.662	0.90	0.10	0.730

Table 9: Simulated data example, Bivariate normal distribution, Scenario 3, $n_0 = n_1 = 100$

References

[1] Pepe MS, Thompson ML. Combining diagnostic test results to increase accuracy. *Biostatistics* 2000; **1**(2):123–140.

[2] Samawi HM, Yin J, Rochani H, Panchal V. Notes on the overlap measure as an alternative to the youden index: How are they related? *Statistics in Medicine* 2017; **36**(26):4230–4240.

[3] Franco-Pereira AM, Nakas CT, Reiser B, Carmen Pardo M. Inference on the overlap coefficient: The binormal approach and alternatives. *Statistical Methods in Medical Research* 2021; **30**(12):2672–2684.

[4] Schmid F, Schmidt A. Nonparametric estimation of the coefficient of overlapping—theory and empirical application. *Computational Statistics & Data Analysis* 2006; **50**(6):1583–1596.

[5] Wang D, Tian L. Parametric methods for confidence interval estimation of overlap coefficients. *Computational Statistics & Data Analysis* 2017; **106**:12–26.

[6] Zhou XH, McClish DK, Obuchowski NA. *Statistical Methods in Diagnostic Medicine*. Wiley-Interscience: New York, 2002.

[7] Hand DJ, Till RJ. A simple generalisation of the area under the roc curve for multiple class classification problems. *Machine Learning* 2001; **45**(2):171–186.

[8] Komaba A, Johno H, Nakamoto K. A novel statistical approach for two-sample testing based on the overlap coefficient 2023.

[9] Su J, Liu J. Linear combinations of multiple diagnostic markers. *Journal of the American Statistical Association* 1993; **88**(424):1350–1355.

[10] Pepe M, Cai T, Longton G. Combining predictors for classification using the area under the receiver operating characteristic curve. *Biometrics* 2006; **62**(1):221–229.

[11] Coolen-Maturi T. Predictive inference for best linear combination of biomarkers subject to limits of detection. *Statistics in Medicine* 2017; **36**(18):2844–2874.

[12] Krzanowski WJ, Hand DJ. *ROC Curves for Continuous Data*. Chapman & Hall/CRC: Boca Raton, FL, 2009.

[13] Wieand S, Gail MH, James BR, James KL. A family of nonparametric statistics for comparing diagnostic markers with paired or unpaired data. *Biometrika* 1989; **76**(3):585–592.

[14] Cox LH, Johnson MM, Kafadar K. Exposition of statistical graphics technology. *ASA Statistical Computing Section*, 1982; 55–56.

[15] Li J, Fine JP. Weighted area under the receiver operating characteristic curve and its application to gene selection. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 2010; **59**(4):673–692.

[16] Montoya JA, Figueroa P G, González-Sánchez D. Statistical inference for the weitzman overlapping coefficient in a family of distributions. *Applied Mathematical Modelling* 2019; **71**:558–568.