Bol. Soc. Paran. Mat.
©SPM – E-ISSN-2175-1188
SPM: www.spm.uem.br/bspm

(3s.) **v. 2026 (44) 8** : 1–12.
ISSN-0037-8712
doi:10.5269/bspm.83886

Bias-Corrected Partial AUC Estimation in Exponential ROC Analysis with Clinical Case Studies

Danisiri Tanuja and Siva G.

ABSTRACT: In diagnostic research, ROC-based performance evaluation is widely used to quantify the discriminative ability of clinical markers. In many practical situations, investigators are interested in a limited false positive range rather than the full curve, which makes the partial area under the ROC curve (pAUC) a more meaningful measure. Existing literature indicates that most pAUC-based developments assume normality, whereas several clinical measurements exhibit non-normal distributions, particularly exponential-type distributions. Under such conditions, conventional pAUC estimates tend to produce biased results and reduce the reliability of interpretation. In this work, we develop a new bias-corrected estimator for partial AUC under the exponential ROC framework. A closed-form expression for the corrected estimator is established, and its statistical behaviour is studied through simulation experiments. Performance is evaluated in terms of bias and mean squared error (MSE). The results demonstrate noticeable improvement over the standard estimator, particularly under non-normal exponential scenarios. The usefulness of the approach is further demonstrated using real clinical case studies. This contribution supports more accurate pAUC computation in exponential ROC settings and can enhance evidence-based medical decision-making.

Keywords: Partial AUC, exponential ROC model, bias-corrected estimation, non-normal distributions, clinical applications.

Contents

1. Introduction

Accurate classification of individuals into clinically relevant groups is a fundamental objective in many statistical and biomedical applications. Receiver Operating Characteristic (ROC) analysis provides a principled framework for assessing the discriminative ability of diagnostic markers and classification models by summarising the trade-off between the false positive rate (FPR = 1 - specificity) and the true positive rate (TPR = sensitivity) across possible decision thresholds. Through this curve, practitioners obtain a global view of test performance and can compare competing markers or models in a standardized manner. ROC estimation can be carried out using either parametric or nonparametric approaches. Parametric, bi-distributional models—most notably the bi-normal model—are widely used due to their interpretability and analytic convenience; see, for example, Egan [1] and the review by Balaswamy and Vishnu Vardhan [2]. However, empirical biomarker measurements frequently depart from normality, exhibiting skewness or heavier tails. To capture such features, researchers have introduced non-normal ROC models based on alternative distributional families, including bi-exponential [3], bi-Lomax [4], and bi-gamma [5] specifications. These flexible parametric forms better accommodate asymmetric and skewed data commonly encountered in practice. Recently, Attwood *et al.,* [6] introduced a skew exponential

2020 *Mathematics Subject Classification:* 62H30, 62P10.
Submitted May 01, 2026.. Published June 05, 2026.

power (SEP) ROC model that reduces cut-point bias and improves estimation efficiency in the presence of heavy-tailed or non-normal data.

A primary scalar summary derived from the ROC curve is the Area under the Curve (AUC). Interpreted probabilistically as $\theta = P(Y > X)$, where X and Y denote marker values from healthy and diseased populations, respectively, the AUC quantifies the probability that a randomly selected diseased subject attains a higher test value than a randomly selected healthy subject [7]. Extensive work exists on estimation and inference for AUC under a range of model assumptions [8,9,10]. Although the AUC is convenient as a global discriminator, clinical decision-making often prioritizes performance within a limited and clinically meaningful range of false positive rates rather than over the entire domain. In such settings, the Partial Area Under the Curve (pAUC) provides a focused measure by restricting integration to a specified FPR interval (c_1, c_2). Early treatments of pAUC and its statistical properties can be found in [11], who developed tests and estimators under normality assumptions for both continuous and ordinal data. Subsequent works extended nonparametric and weighted AUC methods to compare medical imaging modalities and longitudinal markers, treating pAUC as a special case of weighted AUCs [12].

Measurement error (ME) is another pervasive concern in medical studies. Sources of ME include device miscalibration, intra-operator variability, biological fluctuation, and environmental influences; classical examples include blood pressure measurements in children, where observed readings may deviate from true values despite standardized procedures [13]. Measurement error, if unaccounted for, can bias estimates of discriminatory accuracy and lead to misleading conclusions. The statistical literature has proposed various remedies, including reliability and replication-based corrections [14,15], maximum likelihood and replicate-measure approaches under normal error models [16,17], and bias-corrected AUC estimators for both parametric and non-parametric settings [18,19,20]. More recent contributions have addressed complex settings such as mixture and multivariate ROC frameworks with explicit ME considerations [21,22]. Later, Wang and Feng [23] proposed a skew-normal ROC framework to correct diagnostic accuracy measures in the presence of measurement error, particularly for skewed biomarker distributions.

Despite these advances, two significant gaps remain. First, much of the existing ME-focused literature concentrates on the complete AUC and predominantly assumes normality for either the underlying marker distributions or the measurement errors. Second, methodological development for pAUC under non-normal ROC models—particularly exponential-type models that frequently suit skewed clinical measurements—is limited. In practice, clinicians often require inference on pAUC in settings where the biomarker distribution is better represented by an exponential family member and where ME cannot be ignored.

This article addresses these gaps by proposing an approximate bias-corrected estimator for the pAUC within the exponential ROC framework in the presence of measurement error. The estimator is derived analytically, and its finite-sample behaviour is investigated via extensive Monte Carlo simulations spanning a range of sample sizes and error magnitudes. Performance is assessed using bias and mean squared error (MSE) as primary metrics, and comparisons are drawn with the conventional uncorrected pAUC estimator to demonstrate improvements. To illustrate the practical utility of the methodology, it is applied to a real clinical dataset, highlighting how the bias-corrected pAUC can alter interpretation and improve evidence-based decision-making when measurements are noisy.

The remainder of the paper is organized as follows. Section 2 introduces the exponential ROC model, formalises the measurement error structure, and presents the bias-corrected pAUC estimator. Section 3 describes the Monte Carlo simulation design and reports numerical results. Section 4 applies the proposed method to a clinical data example. Section 5 concludes with remarks on limitations, practical guidance, and directions for future research.

2. Proposed Methodology

2.1. Bias Corrected approximation for partial AUC under MEs

Let $X \sim Exp(\lambda_X)$ and $Y \sim Exp(\lambda_Y)$, where λ_X and λ_Y denote the rate parameters corresponding to H and D populations, respectively. The intrinsic measures of the Exponential ROC curve (EROC) are

defined as

$$FPR = e^{-\lambda_X t} \tag{2.1}$$

$$TPR = e^{-\lambda_Y t} \tag{2.2}$$

Based on Equation (2.1), the cut-off value t can be expressed as $t = \frac{-log(FPR)}{\lambda_X}$. Substituting this expression into Equation (2.2) and simplifying yields the corresponding form of the EROC curve as follows

$$\text{EROC (t)} = FPR^{\left(\frac{\lambda_Y}{\lambda_X}\right)} \tag{2.3}$$

If we want to look at the area under the ROC curve from some point c_1 to the point c_2, then the partial area under EROC curve is derived as

$$pAUC = \theta_{\text{p}} = \int_{c_1}^{c_2} EROC(t)\, dt = \frac{c_2^{a+1} - c_1^{a+1}}{a+1}$$

where $a = \frac{\lambda_Y}{\lambda_X}$. When $c_1 = 0$ and $c_2 = 1$, the partial AUC reduces exactly to the full AUC as $\theta_{\text{F}} = \frac{1}{a+1} = \frac{\lambda_X}{\lambda_X+\lambda_Y}$.Using the maximum likelihood estimates of the parameters of λ_X and λ_Y, the natural estimator of θ_{p} is

$$\hat{\theta}_{\text{p}} = \frac{c_2^{\frac{\hat{\lambda}_Y}{\hat{\lambda}_X}+1} - c_1^{\frac{\hat{\lambda}_Y}{\hat{\lambda}_X}+1}}{\frac{\hat{\lambda}_Y}{\hat{\lambda}_X}+1} \tag{2.4}$$

With the help of Taylor series expansion, it can be easily shown that $E(\hat{\theta}_{\text{p}}) = \theta_{\text{p}} + O(n^{*^{-1}})$; where $n^* = min(m, n)$, here m and n are the number of samples in X and Y populations, respectively. If there are MEs present in the observations, it becomes difficult to determine the true values of X and Y accurately. Consequently, the estimated parameters may lack reliability, hindering the ability to proceed with classifying the individuals. To denote these variables that are affected by MEs, we establish the following notation as

$$X_i \text{ observed with } u_i \to A_i, \quad i = 1, 2, \ldots, m,$$
$$Y_j \text{ observed with } v_j \to B_j, \quad j = 1, 2, \ldots, n,$$

where A_i and B_j are the contaminated observations of H and D populations, respectively. Then the natural estimator of θ_{p} for such contaminated observations, as defined

$$\hat{\theta}_{\text{p}_{ME}} = \frac{c_2^{\frac{\hat{\lambda}_B}{\hat{\lambda}_A}+1} - c_1^{\frac{\hat{\lambda}_B}{\hat{\lambda}_A}+1}}{\frac{\hat{\lambda}_B}{\hat{\lambda}_A}+1} \tag{2.5}$$

here, $\hat{\theta}_{\text{p}_{ME}}$ denotes the estimated pAUC of the EROC curve in the presence of MEs. Its expectation can be expressed as $E(\hat{\theta}_{\text{p}_{ME}}) = \theta_{\text{p}} + O(1)$, where $O(1)$ indicates the order of bias introduced by the contaminated observations. Similarly, the full AUC estimator becomes biased under MEs and is formulated as $\hat{\theta}_{\text{F}_{ME}} = \frac{\hat{\lambda}_A}{\hat{\lambda}_A+\hat{\lambda}_B}$. Consequently, both $\hat{\theta}_{\text{p}_{ME}}$ and $\hat{\theta}_{\text{F}_{ME}}$ deviate from their true values, resulting in biased estimates of diagnostic accuracy.
In order to correct the bias in $\hat{\theta}_{\text{p}_{ME}}$,

$$\begin{aligned} E(\hat{\theta}_{\text{p}_{ME}}) &\cong P(B_j > A_i) \\ &\cong P(Y > X + \delta) \\ &\cong \iint [1 - G_Y(s+t)] f_X(s) f_\delta(t) dt ds \end{aligned} \tag{2.6}$$

where, $\delta = (u_i - v_j) \sim N(0, \sigma_u^2 + \sigma_v^2)$, $G_Y(.)$ is the cumulative distribution function of Y and $f_\delta(.)$ is the density function of δ. The integral part in Equation (2.6) cannot be calculated directly since it involves both s and t. So, using the standard Taylor series expansion, we divide this term into two parts, and the expansion term of $G_Y(s+t)$ around $t = 0$ is

$$G_Y(s+t) = G_Y(s) + G'_Y(s)t + G''_Y(s)\frac{t^2}{2!} + G'''_Y(s)\frac{t^3}{3!} + G''''_Y(s + a(t))\frac{t^4}{4!}, \tag{2.7}$$

where $a(t) \in (0, t)$ (see, e.g., Rudin [24]). Then, substituting Equation (2.7) into Equation (2.6), we obtain

$$\begin{aligned} E(\hat{\theta}_{\mathrm{p}_{ME}}) &\cong \iint \left[1 - \left(G_Y(s) + G'_Y(s)\,t + G''_Y(s)\,\tfrac{t^2}{2!} + G'''_Y(s)\,\tfrac{t^3}{3!} + G''''_Y(s + a(t))\,\tfrac{t^4}{4!}\right)\right] f_X(s)\, f_\delta(t)\, dt\, ds \\ &\cong P(Y > X) - \tfrac{E[\delta^2]}{2} E_X[G''_Y(s)] - \iint G''''_Y(s + a(t))\,\tfrac{t^4}{4!}\, f_X(s)\, f_\delta(t)\, dt\, ds \end{aligned}$$

Now the bias can be estimated using $\frac{E[\delta^2]}{2} E_X[G''_Y(s)]$ provided that the last term in the expansion is sufficiently small to be neglected. Since this term is independent of n, asymptotic arguments based on increasing sample size are not applicable in this context. Instead, we will look for conditions on $G_Y(.)$ that will ensure the last term in the expansion is small. The last term will be small enough to be disregarded if the distribution of Y is well approximated by its third-order Taylor expansion, implying that the density of Y behaves approximately like a quadratic function. To formalize this, assume that $E(u^4) = E(v^4) = \nu^2 < \infty$, so that $E(\delta^4) = 2\nu^2 + 6\sigma^4$. Also, suppose $|G_Y^{(4)}(s)| < M$ for $s \in \mathbb{R}$, i.e., that for some small $\zeta > 0$, $\nu^2 + 3\sigma^4 < \frac{12\zeta}{M}$, so $\left|\mathbb{E}(\hat{\theta}_{\mathrm{p}_{ME}}) - P(Y > X) + \sigma^2 \mathbb{E}_X\left[G''_Y(s)\right]\right| \leq M\left(\frac{\mathbb{E}(\delta^4)}{4!}\right) < \zeta$. Then, if ζ is small enough to ignore, $E(\hat{\theta}_{\mathrm{p}_{ME}}) \cong P(Y > X) - B_{\mathrm{p}}$, so $\theta_{\mathrm{p}} = P(Y > X) \cong E(\hat{\theta}_{\mathrm{p}_{ME}}) + B_{\mathrm{p}}$ and the bias in using $\hat{\theta}_{\mathrm{p}_{ME}}$ to estimate $P(Y > X)$ is approximately

$$\begin{aligned} B_{\mathrm{p}} &= \frac{Var(\delta)}{2} \int g'_Y(s) f_X(s) ds \\ &= \frac{\sigma_u^2 + \sigma_v^2}{2} \int_{c_1}^{c_2} \left(-\lambda_Y^2 e^{-\lambda_Y s}\right)\left(\lambda_X e^{-\lambda_X s}\right) ds \qquad (2.8) \\ &= \frac{\sigma_u^2 + \sigma_v^2}{2}(-\lambda_X \lambda_Y^2) \int_{c_1}^{c_2} e^{-(\lambda_X + \lambda_Y)s}\, ds \\ &= \frac{\sigma_u^2 + \sigma_v^2}{2}(-\lambda_X \lambda_Y^2) \left[-\frac{1}{\lambda_X + \lambda_Y} e^{-(\lambda_X + \lambda_Y)s}\right]_{c_1}^{c_2} \\ &= \frac{\sigma_u^2 + \sigma_v^2}{2} \frac{\lambda_X \lambda_Y^2}{\lambda_X + \lambda_Y} \left(e^{-(\lambda_X + \lambda_Y)c_2} - e^{-(\lambda_X + \lambda_Y)c_1}\right) \qquad (2.9) \end{aligned}$$

here, B_{p} represents the bias introduced in estimating $\hat{\theta}_{\mathrm{p}_{ME}}$ due to the presence of MEs. Therefore, the bias-corrected approximation for the partial AUC is

$$\tilde{\theta}_{\mathrm{p}} = \hat{\theta}_{\mathrm{p}_{ME}} + \hat{B}_{\mathrm{p}} \tag{2.10}$$

Measurement error introduces random deviations between the observed and true biomarker values, which affects the estimation of diagnostic accuracy. In the exponential ROC framework, the partial AUC represents the probability that a diseased individual has a higher biomarker value than a healthy individual within a specified false positive range. When measurement error is present, the observed biomarker values become contaminated, leading to incorrect estimation of the distribution parameters and consequently biased pAUC estimates. The proposed bias-correction method quantifies the magnitude of this distortion using a Taylor series approximation and adjusts the contaminated estimator accordingly. This correction effectively compensates for the variability introduced by measurement error and restores the pAUC estimate closer to its true value. As a result, the corrected estimator provides a more reliable assessment of diagnostic performance, particularly in clinical settings where measurement error is unavoidable.

Similarly, the bias expression for estimating the full AUC can be obtained by integrating Equation (2.8) over the entire ROC domain ($c_1 = 0$ to $c_2 = 1$). After simplification, the resulting expression yields the full-AUC bias, denoted as $\hat{B}_{\mathrm{F}}$. The corresponding bias-corrected estimator is then given by

$$\tilde{\theta}_{\mathrm{F}} = \hat{\theta}_{\mathrm{F}_{ME}} + \hat{B}_{\mathrm{F}}$$

The pAUC quantifies classifier performance within a specific region of the ROC curve, often near the origin, where clinically relevant FPRs are small. Since this restricted ROC region may span only a narrow interval, the resulting pAUC values tend to be smaller in magnitude than the full AUC, which can complicate direct interpretation. To place partial area values on a comparable scale, a standardized pAUC measure is employed

$$pAUC_{st} = \frac{1}{2}\left[1 + \frac{pAUC - pAUC_{min}}{pAUC_{max} - pAUC_{min}}\right]$$

where $pAUC_{min} = \frac{(c_2 - c_1)(c_2 + c_1)}{2}$ and $pAUC_{max} = (c_2 - c_1)$. With this transformation, the standardized partial AUC lies between 0.5 (no discrimination) to 1.0 (perfect discrimination), allowing the pAUC to be interpreted on the same scale as the full AUC.

3. Numerical Illustration

3.1. Simulation study

An extensive Monte Carlo simulation study was carried out to assess the performance of the proposed bias-corrected estimator for the EROC curve under different scenarios. The study is divided into two distinct cases: Case 1 represents well-separated populations with a high level of classification accuracy (best classification), whereas Case 2 represents closely overlapping populations that pose a challenging discrimination problem (worst classification). For each case, random samples were generated from an exponential distribution at varying sample sizes and ME levels, reflecting small, moderate, large contamination levels, to examine the impact of MEs on diagnostic accuracy. Both the full area (AUC) and the partial area (pAUC) were estimated for each setting, and the results are presented and discussed below.

Case 1: Best Classification Scenario

In this setting, random samples were generated from exponential distributions i.e., $X \sim Exp(1.8)$ and $Y \sim Exp(0.2)$ for different sample sizes $n = \{25, 50, 100, 200, 500\}$. To mimic the existence of MEs, random error observations were generated at three levels $\sigma_u = \sigma_v = (0.5, 1.0, 2.0)$, and incorporated into the original data to evaluate the performance of the proposed methodology. The results of both the full and partial AUC estimations are summarized in Tables (1) and (2).

From Table (1), it is observed that the estimated AUC values obtained from contaminated data fall below the true AUC, indicating that MEs lead to underestimation of the classifier's discriminative ability. In contrast, the bias-corrected estimators consistently produced values closer to the actual AUC, exhibiting lower bias and mean squared error (MSE) across all sample sizes and error levels compared to the contaminated ones.

A similar trend is evident for the partial area evaluated within the interval $[c_1, c_2] = [0.1, 0.3]$. The corrected estimates demonstrate stability and closeness to the true pAUC, validating the effectiveness of the partial bias-correction procedure in recovering diagnostic accuracy even under contamination. The corresponding partial EROC curves, shown in Figure (1), visually depict the deviation caused by MEs. As the level of contamination increases, the curves for contaminated data shift downward from the true EROC.

Case 2: Worst Classification Scenario This scenario represents a challenging classification problem due to closely overlapping distributions. Random samples are generated with equal parameters, i.e. $X \sim Exp(1.8)$ and $Y \sim Exp(1.8)$ at varying sample sizes n= {20, 50, 100, 200, 500}. As in case (1), error random observations were generated and incorporated into the original data. The corresponding results of full and partial AUCs are presented in Tables (3) and (4), respectively.

Unlike the well-separated case, the estimated AUCs of full and partial are overestimated. For instance, in the partial area at a sample size $n = 500$ under ME level 0.5, the estimated pAUC is 0.64078 which

Table 1: Comparison of contaminated and bias-corrected AUC estimates with associated bias and MSE under Case (1)

$\hat{\theta}_{\mathrm{F}}$	$\sigma_u = \sigma_v$	n	$\hat{\theta}_{\mathrm{F}_{ME}}$	bias	MSE	$\tilde{\theta}_{\mathrm{F}}$	bias	MSE
		25	0.83694	-0.04251	0.00655	**0.88641**	0.00696	0.00028
		50	0.81571	-0.06374	0.00583	**0.88214**	0.00269	0.00021
	0.5	100	0.79856	-0.08089	0.00442	**0.88001**	0.00056	0.00013
		200	0.77894	-0.10051	0.00401	**0.87999**	0.00054	0.00001
		500	0.76832	-0.11113	0.00330	**0.87981**	0.00036	0.00000
		25	0.81975	-0.05970	0.00851	**0.89145**	0.01200	0.00084
		50	0.81054	-0.06891	0.00785	**0.89001**	0.01056	0.00059
0.87945	1.0	100	0.80562	-0.07383	0.00651	**0.88721**	0.00776	0.00037
		200	0.79852	-0.08093	0.00551	**0.88365**	0.00420	0.00025
		500	0.78125	-0.09820	0.00403	**0.88047**	0.00102	0.00008
		25	0.80511	-0.07434	0.00974	**0.89789**	0.01844	0.00098
		50	0.79115	-0.08830	0.00825	**0.89562**	0.01617	0.00068
	2.0	100	0.78932	-0.09013	0.00741	**0.88451**	0.00506	0.00031
		200	0.77452	-0.10493	0.00608	**0.88245**	0.00300	0.00019
		500	0.76541	-0.11404	0.00395	**0.88036**	0.00091	0.00003

$\hat{\theta}_{\mathrm{F}}$ - Actual AUC; $\hat{\theta}_{\mathrm{F}_{ME}}$ - Estimated AUC with ME; $\tilde{\theta}_{\mathrm{F}}$ - Corrected AUC

Table 2: Comparison of true, contaminated, and bias-corrected pAUC estimates with associated bias and MSE under Case (1)

$\hat{\theta}_{\mathrm{p}}$	$\sigma_u = \sigma_v$	n	$\hat{\theta}_{\mathrm{p}_{ME}}$	bias	MSE	$\tilde{\theta}_{\mathrm{p}}$	bias	MSE
		25	0.68772	-0.02651	0.00655	**0.72916**	0.01493	0.00032
		50	0.67011	-0.04412	0.00583	**0.72014**	0.00591	0.00027
	0.5	100	0.66502	-0.04921	0.00442	**0.72001**	0.00578	0.00019
		200	0.66045	-0.05378	0.00401	**0.71804**	0.00381	0.00012
		500	0.65904	-0.05519	0.00330	**0.71560**	0.00137	0.00000
		25	0.67452	-0.03971	0.00851	**0.72999**	0.01576	0.00040
		50	0.66251	-0.05172	0.00785	**0.72875**	0.01452	0.00034
0.71423	1.0	100	0.64259	-0.07164	0.00651	**0.72458**	0.01035	0.00023
		200	0.64017	-0.07406	0.00551	**0.71904**	0.00481	0.00016
		500	0.64999	-0.06424	0.00403	**0.71700**	0.00277	0.00001
		25	0.66354	-0.05069	0.00974	**0.73018**	0.01595	0.00046
		50	0.66028	-0.05395	0.00825	**0.72924**	0.01501	0.00036
	2.0	100	0.65871	-0.05552	0.00741	**0.72355**	0.00932	0.00015
		200	0.65111	-0.06312	0.00608	**0.71998**	0.00575	0.00011
		500	0.64853	-0.06570	0.00395	**0.71606**	0.00183	0.00001

$\hat{\theta}_{\mathrm{p}}$ - Actual AUC; $\hat{\theta}_{\mathrm{p}_{ME}}$ - Estimated AUC with ME; $\tilde{\theta}_{\mathrm{p}}$ - Corrected AUC

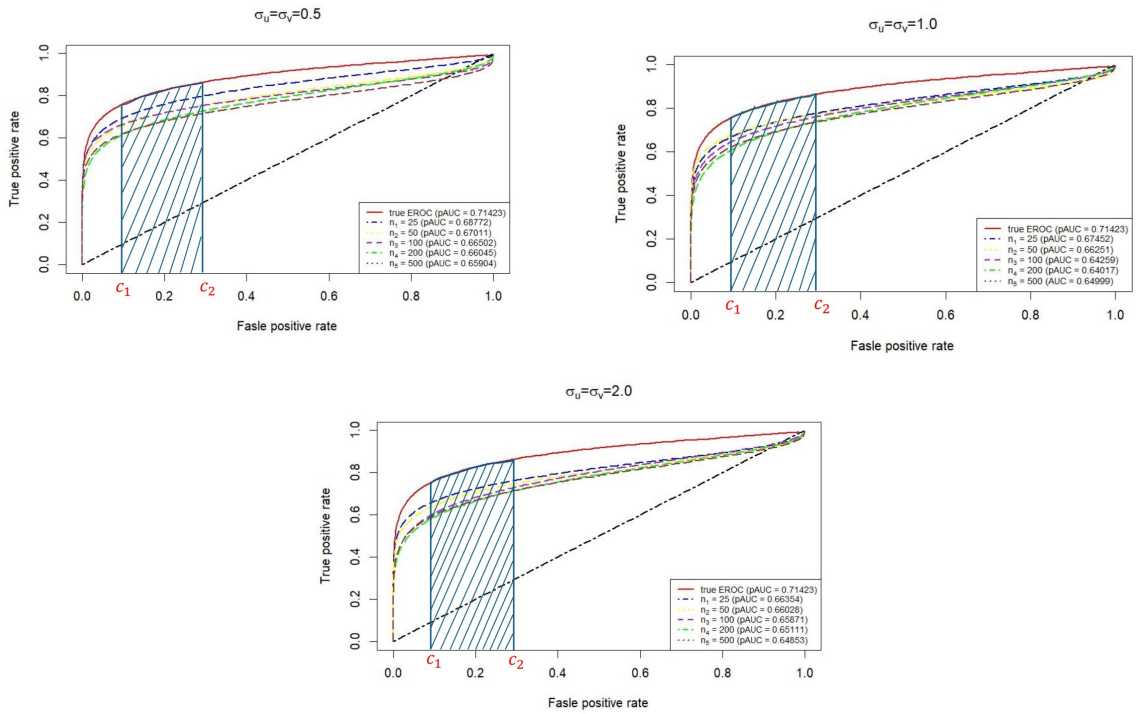

Figure 1: Comparison of EROC curves at different sample sizes under varying levels of MEs case (1)

Table 3: Comparison of contaminated and bias-corrected AUC estimates with associated bias and MSE under Case (2)

$\hat{\theta}_{\mathrm{F}}$	$\sigma_u = \sigma_v$	n	$\hat{\theta}_{\mathrm{F}_{ME}}$	bias	MSE	$\tilde{\theta}_{\mathrm{F}}$	bias	MSE
0.54953	0.5	25	0.61852	0.06899	0.00886	**0.56874**	0.01921	0.00047
		50	0.60154	0.05201	0.00625	**0.55978**	0.01025	0.00032
		100	0.59325	0.04372	0.00531	**0.55721**	0.00768	0.00025
		200	0.58824	0.03871	0.00420	**0.55021**	0.00068	0.00012
		500	0.57564	0.02611	0.00284	**0.54992**	0.00039	0.00001
	1.0	25	0.62965	0.08012	0.01181	**0.56874**	0.01921	0.00287
		50	0.61752	0.06799	0.00965	**0.56021**	0.01068	0.00188
		100	0.60997	0.06044	0.00825	**0.55741**	0.00788	0.00093
		200	0.59475	0.04522	0.00542	**0.55124**	0.00171	0.00037
		500	0.59371	0.04418	0.00334	**0.55003**	0.00050	0.00011
	2.0	25	0.63924	0.08971	0.01461	**0.57125**	0.02172	0.00645
		50	0.61182	0.06229	0.00895	**0.57000**	0.02047	0.00484
		100	0.60965	0.06012	0.00703	**0.56784**	0.01831	0.00301
		200	0.60354	0.05401	0.00482	**0.56124**	0.01171	0.00151
		500	0.59761	0.04808	0.00381	**0.55784**	0.00831	0.00041

$\hat{\theta}_{\mathrm{F}}$ - Actual AUC; $\hat{\theta}_{\mathrm{F}_{ME}}$ - Estimated AUC with ME; $\tilde{\theta}_{\mathrm{F}}$ - Corrected AUC

Table 4: Comparison of true, contaminated, and bias-corrected pAUC estimates with associated bias and MSE under Case (2)

$\hat{\theta}_{\mathrm{p}}$	$\sigma_u = \sigma_v$	n	$\hat{\theta}_{\mathrm{p}_{ME}}$	bias	MSE	$\tilde{\theta}_{\mathrm{p}}$	bias	MSE
0.59143	0.5	25	0.67612	0.08469	0.00886	**0.62904**	0.03761	0.00232
		50	0.66002	0.06859	0.00625	**0.61811**	0.02668	0.00139
		100	0.65893	0.06750	0.00531	**0.61014**	0.01871	0.00089
		200	0.65029	0.05886	0.00420	**0.60079**	0.00936	0.00034
		500	0.64078	0.04935	0.00284	**0.59908**	0.00765	0.00001
	1.0	25	0.69654	0.10511	0.01181	**0.62998**	0.03855	0.00257
		50	0.68942	0.09799	0.00965	**0.62048**	0.02905	0.00152
		100	0.67018	0.07875	0.00825	**0.61961**	0.02818	0.00093
		200	0.66782	0.07639	0.00542	**0.61198**	0.02055	0.00041
		500	0.65781	0.06638	0.00334	**0.59914**	0.00771	0.00001
	2.0	25	0.70589	0.11446	0.01461	**0.63145**	0.04002	0.00302
		50	0.68455	0.09312	0.00895	**0.62547**	0.03404	0.00187
		100	0.66892	0.07749	0.00703	**0.60218**	0.01075	0.00099
		200	0.66025	0.06882	0.00482	**0.60066**	0.00923	0.00031
		500	0.65875	0.06732	0.00381	**0.59508**	0.00365	0.00000

$\hat{\theta}_{\mathrm{p}}$ - Actual AUC; $\hat{\theta}_{\mathrm{p}_{ME}}$ - Estimated AUC with ME; $\tilde{\theta}_{\mathrm{p}}$ - Corrected AUC

is overestimated whereas the bias-corrected pAUC (0.59908) is closely aligned with the actual pAUC (0.59143). It is also observed that as the sample size increases, the proposed bias-corrected estimator pAUC values are getting closer to the actual AUC with negligible error. The graphical representation of the true EROC and EROC (Contaminated) curves are presented in Figure (2). These graphs depict that the actual performance of the classifier gets affected due to the ME. Once, we correct the EROC curve using bias corrected estimator, then the accuracy will be restored. The observed bias behavior differs between the best and worst classification scenarios due to the relative separation between the healthy and diseased distributions. In the best classification scenario, where the distributions are well separated, measurement error introduces additional random variability that reduces the distinguishability between the two groups. This leads to an underestimation of the true AUC, resulting in negative bias. In contrast, in the worst classification scenario, where the distributions substantially overlap, measurement error may artificially increase the apparent separation between observations due to random fluctuations. This can lead to an overestimation of the true AUC, resulting in positive bias. Therefore, the direction and magnitude of bias depend on the degree of overlap between the underlying distributions, highlighting the importance of bias correction, especially in scenarios with measurement uncertainty.

4. Real Dataset

In this section, we discussed the practical utility of the proposed bias-corrected approximation using real datasets. Three real datasets were analyzed: the Indian Liver Patient Dataset (ILPD), Liver Disorders Dataset (LDD), and the Breast Cancer Wisconsin (Diagnostic) Dataset (BCWD). These datasets represent different medical scenarios where MEs are common and illustrate the practical applicability of the bias-corrected estimator in diagnostic tests. In each case, the AUCs (full & Partial) are calculated, and the results were discussed below.

Dataset:1 ILPD [25]

The Indian Liver Patient Dataset (ILPD) consists of 583 records, including 416 individuals diagnosed with liver disease and 167 healthy controls. The present analysis focuses on the biomarker Albumin, a key clinical indicator of hepatic function. The concentration of Albumin in the bloodstream serves as a direct measure of the liver's synthetic capacity, making it a reliable marker for disease discrimination in this study.

Dataset:2 LDD [26]

The Liver Disorders Dataset (LDD) comprises 345 observations with six clinical attributes related to

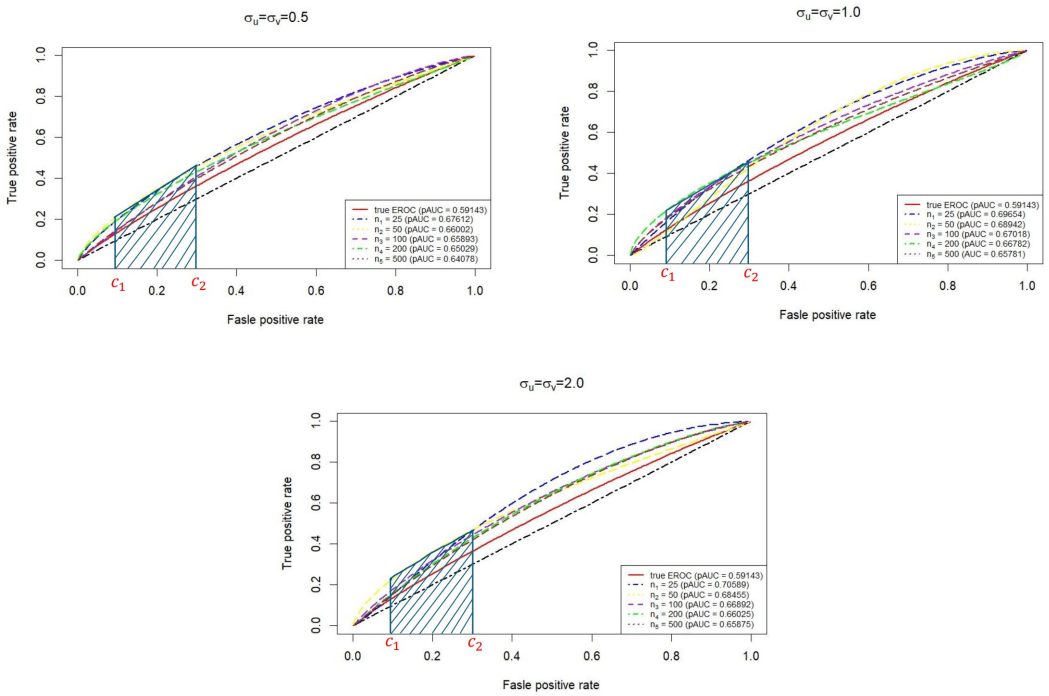

Figure 2: Comparison of EROC curves at different sample sizes under varying levels of MEs for Case (2)

liver enzyme activity. In this study, the variable Alcoholic drinks was selected as the primary biomarker, as excessive alcohol consumption is a major contributing factor to liver dysfunction.

Dataset:3 BCWD [27] This dataset contains 569 samples characterized by 32 attributes derived from digitized images of fine-needle aspirates (FNA) of breast tissue. For the analysis, the attribute Radius was selected, representing the mean distance from the center of a nucleus to its perimeter. This feature serves as a crucial indicator and plays a significant role in differentiating between benign and malignant tumors.

In real datasets, it is often difficult to find explicit MEs. Most available datasets assume that measurements are precise, making it challenging to analyze the impact of errors on diagnostic accuracy. To address this, we incorporated the MEs into the original data by generating random error observations with ME level 0.5. This approach stimulates the datasets where measurements are subject to random fluctuations or inaccuracies. To check the goodness of fit of the contaminated data we applied the Kolmogorov-Smirnov test for each dataset, and the obtained p-values were presented in Table (5), indicating a good fit of the exponential distribution. In Table (6), estimated results of full and partials AUCs, along with their bias and MSEs are reported.

Table 5: Goodness of fit of the Exponential distribution

Datasets	p-values		Fit of the Distribution
	Healthy	Diseased	
ILPD	0.72231	0.63120	Yes
LDD	0.82611	0.60214	Yes
BCWD	0.54164	0.66201	Yes

The presence of MEs in all three datasets results in a notable attenuation of AUC values, reducing

Table 6: Comparison of AUCs between contaminated and bias-corrected estimator of EROC curve

Datasets	$\hat{\theta}_{\mathrm{F}}$	$\hat{\theta}_{\mathrm{F}_{ME}}$	bias	MSE	$\tilde{\theta}_{\mathrm{F}}$	bias	MSE
ILPD	0.52293	0.46185	-0.06107	0.00373	0.52277	-0.00016	0.000001
LDD	0.51073	0.44418	-0.06891	0.00474	0.50666	-0.00406	0.000016
BCWD	0.58977	0.48468	-0.10508	0.01104	0.58232	-0.00744	0.000055
Partial area under the curve (pAUC) at $c_1 = 0.1$ and $c_2 = 0.3$ (Standardized)							
Datasets	$\hat{\theta}_{\mathrm{F}}$	$\hat{\theta}_{\mathrm{F}_{ME}}$	bias	MSE	$\tilde{\theta}_{\mathrm{F}}$	bias	MSE
ILPD	0.52291	0.46182	-0.06109	0.00373	0.52273	-0.00018	0.000000
LDD	0.51071	0.44416	-0.06655	0.00443	0.50664	-0.00407	0.000017
BCWD	0.58975	0.48466	-0.10509	0.01104	0.58230	-0.00745	0.000056

the discriminative ability of the clinical markers. For example, in the ILPD dataset, the contaminated AUC decreases from 0.52293 to 0.46185, leading to a higher bias (-0.06107) and inflated MSE (0.00373). After applying the proposed bias-corrected approach, the corrected AUC (0.52277) closely aligns with the original AUC, with significantly lower bias and MSE, confirming reliable recovery of the true diagnostic performance. Similar patterns are also observed in the LDD and BCWD datasets, where MEs distort the AUC to a greater extent, and the bias-corrected estimator effectively corrects for this distortion. These results confirm that ignoring MEs can lead to an underestimation of classifier accuracy.

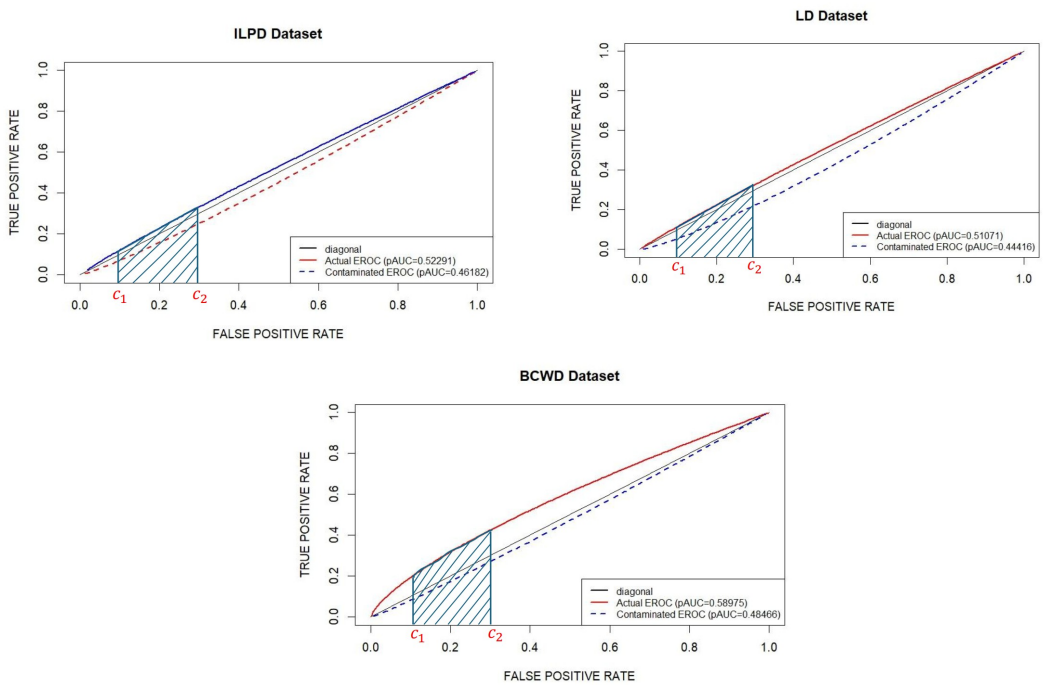

Figure 3: The true and contaminated EROC (partial) curves for real datasets

The pAUC estimates further support these findings. In all three datasets, contaminated pAUC values exhibit substantial downward bias, whereas bias-corrected pAUC estimates restore accuracy with negligible error. This emphasizes the importance of incorporating measurement-error correction when evaluating performance within a specific region of the ROC curve.

From these results, it is understood that the accuracy was affected by due to the error observations involved in the dataset. In such situations, using the bias-corrected estimator, the true accuracy values are achieved which has minimum bias and MSE when compared to the estimated AUC. Figure (3) presents the EROC curves before and after measurement error correction for each dataset. The curves under contamination clearly deteriorate and shift downwards to the diagonal line, indicating a loss in classification ability. Therefore, the empirical findings from these three real-world datasets validate that the bias-corrected estimator reliably mitigates the detrimental effects of MEs, thereby enhancing both global and regional accuracy assessments in diagnostic testing.

5. Conclusion

This study examined the impact of measurement error on ROC-based diagnostic evaluation under non-normal distributional settings. Results showed that measurement error can substantially bias both complete AUC and partial AUC estimations, leading to misinterpretations regarding the true discriminatory capacity of a biomarker. To address this, a bias-corrected estimator was developed within the exponential ROC framework, and its performance was assessed through Monte Carlo simulations and real clinical datasets. The proposed estimator consistently reduced bias while maintaining low MSE, demonstrating improved reliability in both global and region-specific diagnostic assessments.

While the findings confirm the usefulness of the proposed method, certain limitations warrant additional investigation. This work focused primarily on exponential-type non-normal distributions, and future evaluation across broader families of skewed or heavy-tailed distributions would enhance generalisability. Furthermore, the current correction assumed classical additive measurement error; more complex structures, such as heteroscedastic, Berkson-type, and correlated replicate errors, were not addressed. Real clinical settings may also involve missing data, covariate effects, multi-class disease states, or longitudinal follow-up, which were beyond the present scope.

Future research may extend the methodology to semi-parametric and mixture-based ROC models, incorporate multi-marker fusion, and develop inferential tools for variance estimation and interval estimation for pAUC under measurement error. Exploring computationally efficient implementations for high-dimensional data and embedding measurement-error correction within clinical decision support environments also represents an important step toward broader application in routine diagnostic practice.

References

1. Egan, J. P., *Signal detection theory and ROC-analysis*, Academic Press, (1975).
2. Balaswamy, S., and Vardhan, R. V., *An Anthology of Parametric ROC Models*, Res. Rev.: J. Stat. 5(2), 32–46, (2016).
3. England, W. L., *An exponential model used for optimal threshold selection on ROC curves*, Med. Decis. Making 8(2), 120–131, (1988).
4. Campbell, G., and Ratnaparkhi, M. V., *An application of Lomax distributions in receiver operating characteristic (ROC) curve analysis*, Commun. Stat. Theory Methods 22(6), 1681–1687, (1993).
5. Hussain, E., *The bi-gamma ROC curve in a straightforward manner*, J. Basic Appl. Sci. 8, 309–314, (2012).
6. Attwood, K., Hou, S., and Hutson, A., *Application of the skew exponential power distribution to ROC curves*, J. Appl. Stat. 50(8), 1709–1724, (2023).
7. Bamber, D., *The area above the ordinal dominance graph and the area below the receiver operating characteristic graph*, J. Math. Psychol. 12, 387–415, (1975).
8. Hanley, J. A., and McNeil, B. J., *A method of comparing the areas under receiver operating characteristic curves derived from the same cases*, Radiology 148(3), 839–843, (1983).
9. Faraggi, D., and Reiser, B., *Estimation of the area under the ROC curve*, Stat. Med. 21, 3093–3106, (2002).
10. Vardhan, R. V., and Sarma, K. V. S., *Estimation of the area under the ROC curve using confidence intervals of mean*, ANU J. Phys. Sci. 2(1), 29–39, (2010).
11. McClish, D. K., *Analyzing a portion of the ROC curve*, Med. Decis. Making 9(3), 190–195, (1989).
12. Tang, L. L., Liu, A., Chen, Z., Schisterman, E. F., Zhang, B., and Miao, Z., *Nonparametric ROC summary statistics for correlated diagnostic marker data*, Stat. Med. 32(13), 2209–2220, (2013).
13. Shear, C. L., Burke, G. L., Freedman, D. S., Webber, L. S., and Berenson, G. S., *Designation of children with high blood pressure—considerations on percentile cut points and subsequent high blood pressure: the Bogalusa Heart Study*, Am. J. Epidemiol. 125, 73–84, (1987).

14. Dunn, G., *Design and analysis of reliability studies: The statistical evaluation of measurement errors*, Edward Arnold Publishers, (1989).

15. Schisterman, E. F., Faraggi, D., Reiser, B., and Trevisan, M., *Statistical inference for the area under the receiver operating characteristic curve in the presence of random measurement error*, Am. J. Epidemiol. 154(2), 174–179, (2001).

16. Tosteson, T. D., Buonaccorsi, J. P., Demidenko, E., and Wells, W. A., *Measurement error and confidence intervals for ROC curves*, Biom. J. 47(4), 409–416, (2005).

17. Perkins, N. J., Schisterman, E. F., and Vexler, A., *Generalized ROC curve inference for a biomarker subject to a limit of detection and measurement error*, Stat. Med. 28, 1841–1860, (2009).

18. Coffin, M., and Sukhatme, S., *A parametric approach to measurement errors in receiver operating characteristic studies*, Lifetime Data: Models in Reliability and Survival Analysis, 71–75, (1996).

19. Coffin, M., and Sukhatme, S., *Receiver operating characteristic studies and measurement errors*, Biometrics, 823–837, (1997).

20. Kim, J., and Gleser, L. J., *SIMEX approaches to measurement error in ROC studies*, Commun. Stat. Theory Methods 29(11), 2473–2491, (2000).

21. Siva, G., Vardhan, R. V., and Kamath, A., *Estimating the AUC of the MROC curve in the presence of measurement errors*, Commun. Stat. Appl. Methods 29(5), 533–545, (2022).

22. Siva, G., and Vardhan, R. V., *Estimation of AUC of Mixture ROC Curve in the Presence of Measurement Errors*, Stat. Appl. 21(1), 1–12, (2023).

23. Wang, C. Y., and Feng, Z., *A flexible method for diagnostic accuracy with biomarker measurement error*, Mathematics 11(3), 549, (2023).

24. Rudin, W., *Principles of Mathematical Analysis*, 3^{rd} Edition, McGraw-Hill, New York, (1976).

25. Ramana, B., and Venkateswarlu, N., *ILPD (Indian Liver Patient Dataset)*, UCI Machine Learning Repository, `https://doi.org/10.24432/C5D02C`.

26. Forsyth, R. S., *Liver Disorders*, UCI Machine Learning Repository, `https://doi.org/10.24432/C54G67`, (2016).

27. Wolberg, W., Mangasarian, O., Street, N., and Street, W., *Breast Cancer Wisconsin (Diagnostic)*, UCI Machine Learning Repository, `https://doi.org/10.24432/C5DW2B`.

Danisiri Tanuja,
Department of Mathematics,
VIT-AP University, Amaravati,
Andhra Pradesh, India - 522 241
E-mail address: `tanuja.23phd7242@vitap.ac.in`

and

Siva G.,
Department of Mathematics,
VIT-AP University, Amaravati,
Andhra Pradesh, India - 522 241
E-mail address: `siva.g@vitap.ac.in`