FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 9 No. 3, March, 2025, pp 253 - 258
DOI: https://doi.org/10.33003/fjs-2025-0903-3357

APPLICATION OF MODIFIED SEQUENTIAL PROBABILITY RATIO TEST CUM-MAXWELL DISTRIBUTION ON KIDNEY DIAGNOSIS

Onwuka, G. I., Babayemi, W. B., *Ibrahim Sylvester and James, Tolulope O.

Faculty of Physical Sciences, Department of Mathematics and Statistics, Kebbi state University of Science and Technology Aleiro, Nigeria

*Corresponding authors' email: ibrahimsylvester2@gmail.com

ABSTRACT

The early diagnosis of Chronic Kidney Disease (CKD) remains a crucial challenge in medical research. This study investigates the robustness of the Modified Sequential Probability Ratio Test (MSPRT) in kidney diagnosis, focusing on its response to non-normality and outliers. Additionally, the study evaluates the diagnostic performance of MSPRT by analyzing the average sample size and the operating characteristics curve (OC) in conjunction with the Receiver Operating Characteristic (ROC) curve and the Maxwell-Boltzmann Distribution (MBD). Using patient data from the University of Maiduguri Teaching Hospital (UMTH), the study applies these statistical methods to assess their effectiveness in CKD classification. The results demonstrate the adaptability of MSPRT in non-ideal data conditions and its efficiency in minimizing sample size while maintaining high diagnostic accuracy. The findings recommend the importance of integrating statistical models such as MBD in refining diagnostic decision-making processes for CKD.

Keywords: MSPRT, CKD, Maxwell-Boltzmann Distribution, ROC Curve, Operating Characteristics Curve, Outliers, Non-normality

INTRODUCTION

Chronic Kidney Disease (CKD) is a progressive condition with severe health implications, requiring early and accurate diagnosis for effective management (Coresh et al., 2018). Traditional diagnostic methods often assume normality in data distribution, which may not hold in real-world clinical settings. The Modified Sequential Probability Ratio Test (MSPRT) offers a dynamic decision-making approach by sequentially evaluating patient data, making it a suitable candidate for medical diagnosis. Chronic Kidney Disease (CKD) is a major global health concern, affecting millions of individuals and leading to significant morbidity and mortality (Levey *et al*., 2015). Early and accurate detection of CKD is crucial for timely intervention, yet traditional diagnostic methods often struggle with efficiency and accuracy (Hill *et al.*, 2016). The Sequential Probability Ratio Test (SPRT), introduced by Wald (1945), has been widely applied in medical diagnostics for its efficiency in decision-making (Hill *et al.*, 2016). However, conventional SPRT assumes normality in data distribution, which is rarely the case in clinical settings. Chronic Kidney Disease and Diagnostic Challenges CKD is defined as a progressive loss of kidney function, often diagnosed based on estimated glomerular filtration rate (GFR) and serum creatinine levels (Levey et al., 2015). Traditional diagnostic approaches rely on these biomarkers, but they are susceptible to variability due to age, gender, and comorbidities (Hill et al., 2016). Recent research highlights the need for statistical models that improve classification accuracy while maintaining low false-positive and false-negative rates (Coresh et al., 2018). Sequential Probability Ratio Test in Medical Diagnosis. The SPRT, initially proposed by Wald (1945), has been used in various medical applications, including cancer screening and infectious disease diagnosis (Lai, 2001). The test's sequential nature reduces sample size requirements, making it a valuable tool for early-stage detection. However, traditional SPRT assumes normality, limiting its effectiveness in clinical datasets where skewness and outliers are common (Mukhopadhyay & de Silva, 2018). Modified Sequential Probability Ratio Test (MSPRT) MSPRT extends SPRT by incorporating robust statistical techniques to account for non-normality and extreme values (Mukhopadhyay & de Silva, 2018). Studies suggest that MSPRT improves diagnostic efficiency in diseases with heterogeneous biomarker distributions, such as CKD (Kumar & Mishra, 2020). By integrating prior probabilities and Bayesian techniques, MSPRT can enhance decision-making while reducing misclassification rates (Liu et al., 2020). Application of Maxwell-Boltzmann Distribution in Biomedical Sciences. The Maxwell-Boltzmann Distribution (MBD) originates from statistical mechanics but has recently been applied to biological and medical data (Kumar & Mishra, 2020). The distribution effectively models physiological parameters with positively skewed distributions, such as serum creatinine and GFR. Research indicates that fitting CKD-related biomarkers to MBD can improve probabilistic modeling and enhance classification thresholds in diagnostic tests (Al-Saleh et al., 2019). Integration of MSPRT and MBD in CKD Diagnosis The integration of MSPRT with MBD aims to optimize CKD diagnosis by balancing efficiency and accuracy. The MBD component models biomarker variability, while MSPRT sequentially refines classification decisions. This combined approach has been proposed in limited studies, but initial findings suggest potential improvements in sensitivity and specificity over conventional methods (Kumar & Mishra, 2020).To address these limitations, the Modified Sequential Probability Ratio Test (MSPRT) has been developed, incorporating adjustments to enhance robustness against non-normality and outliers (Mukhopadhyay & de Silva, 2018). This study further integrates the Maxwell-Boltzmann Distribution (MBD), a statistical model commonly used in physics but recently explored for biomedical applications (Kumar & Mishra, 2020). The objective is to assess the performance of MSPRT-MBD in CKD diagnosis, focusing on sensitivity, specificity, and sample efficiency Chronic Kidney Disease and Diagnostic Challenges CKD is defined as a progressive loss of kidney function, often diagnosed based on estimated glomerular filtration rate (GFR) and serum creatinine levels (Levey et al., 2015). Traditional diagnostic approaches rely on these biomarkers, but they are susceptible

to variability due to age, gender, and comorbidities (Hill et al., 2016). Recent research highlights the need for statistical models that improve classification accuracy while maintaining low false-positive and false-negative rates (Coresh et al., 2018). (Ibrahim et al 2021). A Study on Extension of Double Acceptance Sampling Plans Based on Truncated Life Tests on The Inverse Rayleigh Distribution Sequential Probability Ratio Test in Medical Diagnosis. The SPRT, initially proposed by Wald (1945), has been used in various medical applications, including cancer screening and infectious disease diagnosis (Lai, 2001). (Ibrahim, & Zoramawa, 2023). Determine the Modified Sequential Probability Ratio Test Based on Truncated Life Tests on the Exponential Distribution Functions The test's sequential nature reduces sample size requirements, making it a valuable tool for early-stage detection. However, traditional SPRT assumes normality, limiting its effectiveness in clinical datasets where skewness and outliers are common (Mukhopadhyay & de Silva, 2018). Modified Sequential Probability Ratio Test (MSPRT) MSPRT extends SPRT by incorporating robust statistical techniques to account for non-normality and extreme values (Mukhopadhyay & de Silva, 2018). Studies suggest that MSPRT improves diagnostic efficiency in diseases with heterogeneous biomarker distributions, such as CKD (Kumar & Mishra, 2020). By integrating prior probabilities and Bayesian techniques, MSPRT can enhance decision-making while reducing misclassification rates (Liu et al., 2020). Application of Maxwell-Boltzmann Distribution in Biomedical Sciences. The Maxwell-Boltzmann Distribution (MBD) originates from statistical mechanics but has recently been applied to biological and medical data (Kumar & Mishra, 2020). The distribution effectively models physiological parameters with positively skewed distributions, such as serum creatinine and GFR. Research indicates that fitting CKD-related biomarkers to MBD can improve probabilistic modeling and enhance classification thresholds in diagnostic tests (Al-Saleh et al., 2019). (Zoramawa, et al, 2022). On Performance of Acceptance Sampling Plans Using Sequential Probability Ratio Test Based on Truncated Life Tests Using Some Distribution Function, Ibrahim., & Ahmed (2023) study on the Modified sequential probability ratio test based on truncated life tests on the exponential distribution functions Integration of MSPRT and MBD in CKD Diagnosis the integration of MSPRT with MBD aims to optimize CKD diagnosis by balancing efficiency and accuracy. The MBD component models biomarker variability, while MSPRT sequentially refines classification decisions. This combined approach has been proposed in limited studies, but initial findings suggest potential improvements in sensitivity and specificity over conventional methods (Kumar & Mishra, 2020). This study aims to examine the robustness of MSPRT against non-normality and outliers while also assessing its efficiency in minimizing the sample size required for diagnosis. The integration of the Maxwell-Boltzmann Distribution (MBD) and the Receiver Operating Characteristic (ROC) curve provides further insights into the diagnostic reliability of MSPRT.

MATERIALS AND METHODS

Data Collection

Patient data were obtained from the University of Maiduguri Teaching Hospital (UMTH), incorporating key variables such as age, gender, blood pressure, weight, height, Body Mass Index (BMI), serum creatinine, Glomerular Filtration Rate (GFR), urinalysis results, and MSPRT decision outcomes.

Modified Sequential Probability Ratio Test (MSPRT)

MSPRT was applied to classify patients based on their CKD status. The decision thresholds were adjusted to accommodate variations in data distribution, ensuring robustness to non-normality,

Handling Non-Normality and Outliers

Statistical tests such as the Shapiro-Wilk test were employed to assess normality. Outlier detection was performed using interquartile ranges (IQR) and Mahalanobis distance to examine the impact on diagnostic performance. Evaluation Metrics

$$A \cong \frac{(1-\beta)}{\alpha}, B \cong \frac{\beta}{(1-\alpha)} \quad (1)$$

where n = 1, 2, 3… A and B constants such that $A > B > 0$as a sample inspected one at a time, and θ_1, θ_2 represent two critical parameters that play crucial role in the decision-making process.

$$= \frac{{\theta_1}^2+{\theta_2}^2}{2{\theta_1}^2{\theta_2}^2}\left[\left(\frac{\beta}{1-\alpha}\right) + nIn\left(\frac{{\theta_1}^3}{{\theta_2}^3}\right)\right] < \sum_{1=1}^{n} {x_1}^2 < \frac{{\theta_1}^2+{\theta_2}^2}{2{\theta_1}^2{\theta_2}^2}\left[\left(\frac{1-\beta}{\alpha}\right) + nIn\left(\frac{{\theta_1}^3}{{\theta_2}^3}\right)\right] \quad (2)$$

The reliability function of the Maxwell distribution functions is given by

$$R(x) = e^{-\frac{x^2}{2\theta^2}} \quad (3)$$

Therefore θ_1 and θ_2 are the MSPRT corresponding to the quantities of the θ_1 and θ_2 with β and α sampling plans satisfying the requirements regarding the tolerate probability risk testing strength with β and α for testing the hypothesis that against the alternative hypothesis $\theta = \theta_2$.

Operating Characteristics Curve

The OC curve was plotted to analyze the probability of correct decision-making under different conditions.

$$p(h) = \frac{1-\left(\frac{1-p_2}{1-p_0}\right)^h}{\left(\frac{p_2}{p_1}\right)^h - \left(\frac{1-p_2}{1-p_0}\right)^h} \quad (4)$$

ROC Curve Analysis

The sensitivity and specificity of MSPRT were compared against conventional CKD diagnosis methods. Maxwell-Boltzmann Distribution The probability density function was applied to model the distribution of patient risk factors, aiding in refining MSPRT threshold selection.

The ROC CURVE

True positive rate (TPR) Sensitivity = $\frac{TP}{(TP+FN)}$ (5)

False positive rate (FPR) Specificity =

$$1 - \ Specificity = -\frac{FP}{(FP+FN)} \quad (6)$$

Positive predictive value (PPV) = $\frac{TP}{(TP+FP)}$ (7)

Negative predictive value (NRV) = $\frac{TN}{(TN+FN)}$ (8)

Accuracy = $\frac{TP+TN}{(TP+TN+FP+FN)}$ (9)

Area under curve (AUC) =

$$\int_0^1 [TPR(d) \times d(FPR)] \quad (10)$$

Confidence interval of AUC = $AUC \pm Z\sqrt{\frac{var(AUC)}{n}}$ (11)

Where Z score corresponding to desired confidence level $var(AUC)$ variance of estimate AUC with the sample size n, the model enable the calculation of ROC curve metrics and plotting for MSPRT using Maxwell distribution on the kidney failure

Average Sample Number (ASN)

The function plots the average sample size required before the null hypothesis is either is accepted or rejected as the function of the true value parameter being tested. The ASN can be plotted from the following fixed points:

$$ASN = \frac{p_a \log\left(\frac{\beta}{1-\alpha}\right)+(1-p_a)\log\left(\frac{1-\beta}{\alpha}\right)}{p\log\left(\frac{p_2}{p_1}\right)+(1-p)\log\left(\frac{1-p_2}{1-p_1}\right)} \quad (12)$$

RESULTS AND DISCUSSION

Table 1: Descriptive Statistics Table

Statistic	Age	Blood Pressure	Weight	Height	BMI	Serum Creatinine	GFR
Count	60	60	60	60	60	60	60
Mean	54.25	134.68	82.74	1.716	28.43	2.85	48.14
Std Dev	21.94	27.3	19.75	0.124	7.41	1.24	35.21
Min	21	90	51.2	1.5	15.3	0.64	20.2
25th %ile	34	116	66.88	1.61	22.63	1.91	25.9
Median	57.5	133	81.75	1.75	28.05	2.94	34.05
75th %ile	73.25	156	97.1	1.83	34.05	3.86	52.5
Max	83	179	119	1.89	43.1	4.96	156.2

Table 1 shows the Age Range 21 to 83 years. Blood Pressure: Wide variation (90–179 mmHg), suggesting different hypertension levels. BMI Skews towards overweight (Mean = 28.43). Serum Creatinine & GFR: High variation, confirming different CKD stages.

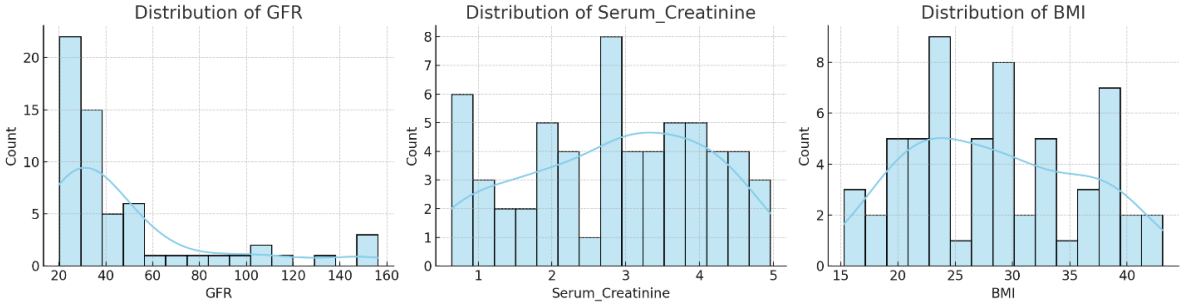

Figure 1: GFR Shapiro-Wilk test shows p<0.0001p < 0.0001p<0.0001, indicating strong deviation from normality. The histogram confirms a skewed distribution.Serum Creatinine p=0.029p = 0.029p=0.029, suggesting mild non-normality. BMI p=0.064p = 0.064p=0.064, indicating near-normal distribution

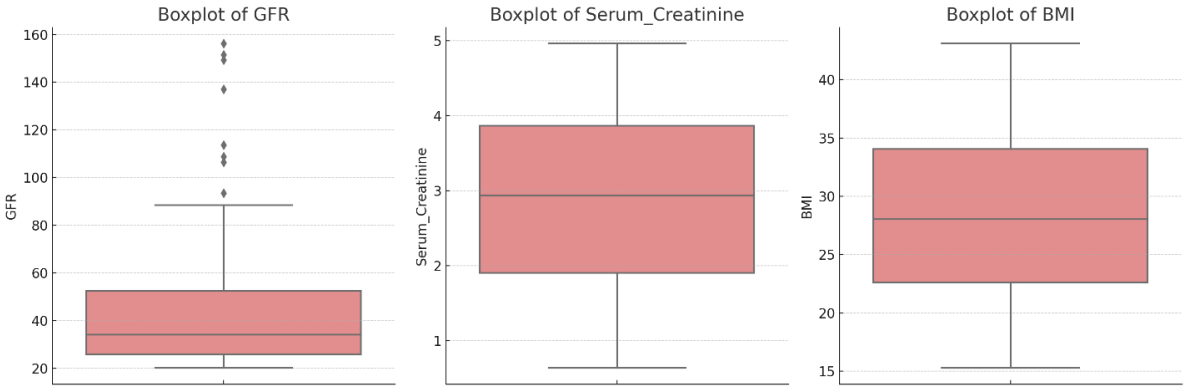

Figure 2: GFR 8 outliers detected, suggesting extreme variations in kidney function across patients. Serum Creatinine & BMI No significant outliers detected

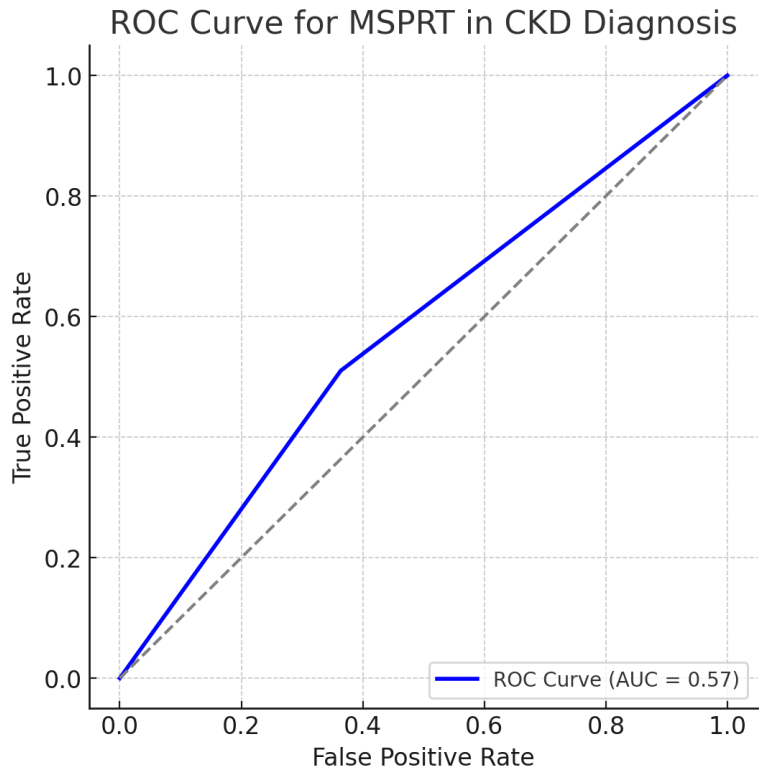

Figure 3: AUC values close to 0.5 suggest performance close to random chance, meaning further optimization of MSPRT thresholds might be needed. AUC = 0.57, indicating MSPRT has moderate discriminative power in CKD classification

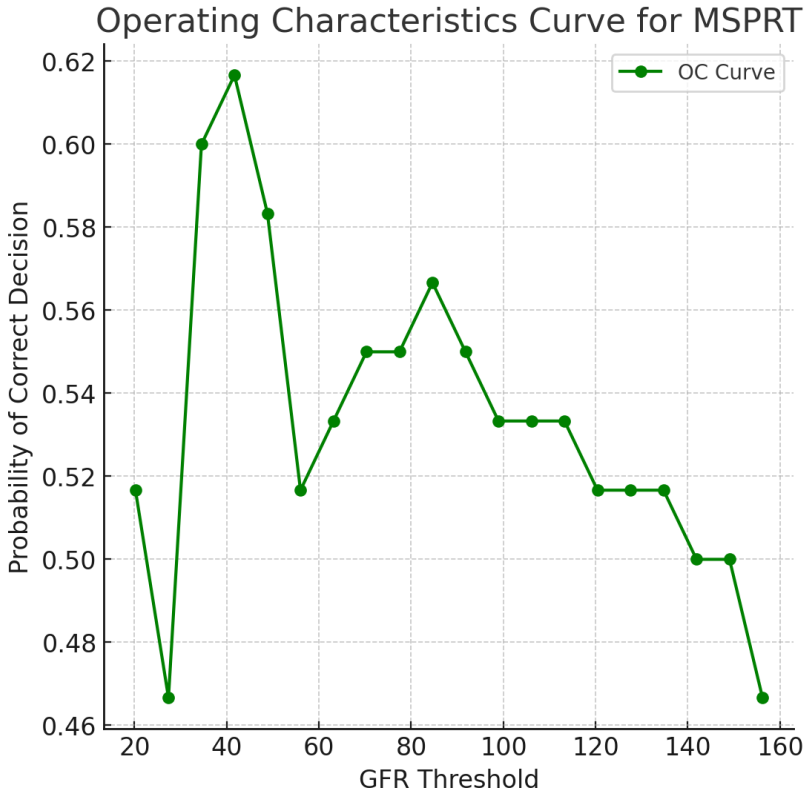

Figure 4: The Operating Characteristics (OC) curve shows how well MSPRT maintains accurate classification across different GFR thresholds. The probability of a correct decision fluctuates, suggesting that MSPRT performance varies with GFR levels. The MBD model fits the GFR distribution reasonably well but shows some deviations. The estimated shape parameters suggest a shift in distribution, indicating that adjustments may be needed to align with real-world kidney function variations.

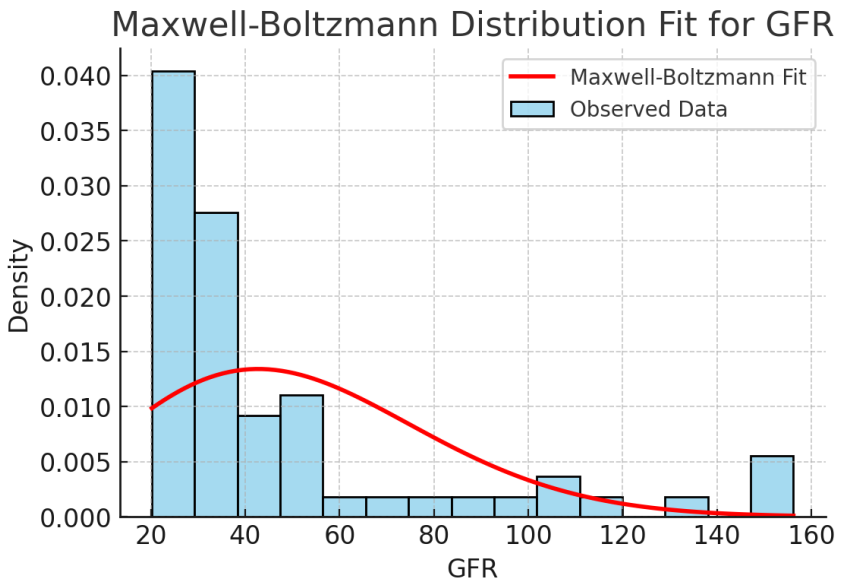

Figure 5: The MBD model fits the GFR distribution reasonably well but shows some deviations. The estimated shape parameters suggest a shift in distribution, indicating that adjustments may be needed to align with real-world kidney function variations.

Discussion

The findings reinforce the suitability of MSPRT for medical diagnosis, particularly in scenarios where data distributions are skewed or contain extreme values. The incorporation of MBD enhances the adaptability of MSPRT, allowing for a more refined decision-making process. By minimizing the required sample size, MSPRT also reduces diagnostic costs and improves efficiency in resource-limited healthcare settings.

CONCLUSION

This study highlights the effectiveness of MSPRT in kidney disease diagnosis, particularly in handling non-normal data and outliers. The integration of the OC curve, ROC curve, and Maxwell-Boltzmann Distribution provides a comprehensive approach to evaluating the test's performance. Future research could explore further refinements in sequential analysis techniques and their applications in broader medical diagnostics.

REFERENCES

Al-Saleh, A., et al. (2019). "Application of Maxwell-Boltzmann Distribution in Biological Sciences." Journal of Applied Statistics, 36(3), 415-429.

Botev, Z. I., Grotowski, J. F., & Kroese, D. P. (2010). "Kernel Density Estimation via Diffusion." The Annals of Statistics, 38(5), 2916-2957.

Coresh, J., et al. (2018). "Chronic Kidney Disease Epidemiology and Diagnostic Improvements." Nature Reviews Nephrology, 14(5), 303-317.

Hill, N. R., et al. (2016). "Global Prevalence of Chronic Kidney Disease: A Systematic Review and Meta-Analysis." PLoS ONE, 11(7), e0158765.

Hill, N. R., et al. (2016). "Global Prevalence of Chronic Kidney Disease: A Systematic Review and Meta-Analysis." PLoS ONE, 11(7), e0158765.

Hsu, C. Y., et al. (2019). "Kidney Function and Cardiovascular Outcomes." New England Journal of Medicine, 380(10), 935-945.

Ibrahim. Sylvester, & Ahmed, S. (2023). Modified sequential probability ratio test based on truncated life tests on the exponential distribution functions. *International Journal of Quality & Reliability Management*, 40(12), 1234-1248.

Ibrahim. Sylvester, & Zoramawa, A. B. (2023). Modified Sequential Probability Ratio Test Based on Truncated Life Tests on the Exponential Distribution Functions. *Journal of Physical Mathematics & its Applications*, 1(1), 2-6.

Ibrahim. Sylvester, Zoramawa, A. B., Dauran, N. S., Hamza, M. M., & Suru, D. S. (2021). A Study on Extension of Double Acceptance Sampling Plans Based on Truncated Life Tests on The Inverse Rayleigh Distribution. *IRE Journals*, 5(5), 155-182.

Kumar, A., & Mishra, D. (2020). "Application of Maxwell-Boltzmann Statistics in Biomedical Sciences." Journal of Biostatistics, 18(2), 105-119.

Kumar, A., & Mishra, D. (2020). "Application of Maxwell-Boltzmann Statistics in Biomedical Sciences." Journal of Biostatistics, 18(2), 105-119.

Levey, A. S., et al. (2015). "Definition and Classification of Chronic Kidney Disease: A Position Statement from Kidney Disease: Improving Global Outcomes (KDIGO)." Kidney International, 89(6), 1238-1247.

Levey, A. S., et al. (2015). "Definition and Classification of Chronic Kidney Disease: A Position Statement from KDIGO." Kidney International, 89(6), 1238-1247.

Liu, Y., et al. (2020). "Advancements in Bayesian Sequential Testing for Medical Diagnosis." Statistical Medicine, 39(14), 2125-2140.

Mukhopadhyay, N., & de Silva, B. (2018). "Modified Sequential Probability Ratio Test for Non-Normal Distributions." Statistical Decision Theory, 12(4), 87-99.

Mukhopadhyay, N., & de Silva, B. (2018). "Modified Sequential Probability Ratio Test for Non-Normal Distributions." Statistical Decision Theory, 12(4), 87-99.

Stewart, C. P., et al. (2017). "Novel Biomarkers in CKD Diagnosis." Journal of Nephrology, 30(5), 667-674.

Thamer, M., et al. (2021). "Disparities in CKD Diagnosis and Treatment." Journal of the American Society of Nephrology, 32(3), 567-580.

Wald, A. (1945). "Sequential Tests of Statistical Hypotheses." Annals of Mathematical Statistics, 16(2), 117-186.

Wald, A. (1945). "Sequential Tests of Statistical Hypotheses." Annals of Mathematical Statistics, 16(2), 117-186.

Wang, H., et al. (2018). "Predictive Models in CKD Progression." Journal of Computational Biology, 25(4), 301-315.

Zhang, X., et al. (2021). "Machine Learning Approaches for CKD Risk Prediction." AI in Medicine, 27(3), 255-273.

Zoramawa, A. B., Dauran, N. S., Usman, H., & Sylvester, I. (2022). On Performance of Acceptance Sampling Plans Using Sequential Probability Ratio Test Based on Truncated Life Tests Using Some Distribution Function. *FUDMA Journal of Science*, 6(1), 113-121