DEPARTMENT OF STATISTICS
AHMADU BELLO UNIVERSITY, ZARIA
2nd International Conference on Statistical Sciences and Data Analytics (ICSSDA 2026) Proceedings Paper

Enhancing Patient-Specific Bayesian Modified Sequential Probability Ratio Test for Personalised Kidney Disease Diagnosis

Ibrahim Sylvester[1] Isaac Amos[2], David Saiki Suru[1], Mustapha Usman[3]

[1,1]Department of Mathematics, Abdullahi Fodiyu University of Science and Technology, Aleiro, Nigeria
[2]Federal Polytechnic Monguno, Department of Statistics, Borno State, Nigeria
[3]Abdulkadir Benishek Model Staff School, Borno State, Kashim Ibrahim University
*Correspondence: ibrahimsylvester2@gmail.com

Abstract

Kidney disease remains a major public health concern because delayed diagnosis can increase the risk of disease progression and adverse outcomes. Conventional diagnostic procedures may rely on fixed decision rules and may not adequately reflect differences in patient risk. This study developed an Enhanced Patient-Specific Bayesian Modified Sequential Probability Ratio Test (PS-BMSPRT) for personalised kidney disease diagnosis by incorporating patient-specific informative prior probabilities into a sequential Bayesian decision framework. The study used the Chronic Kidney Disease dataset cited in the original manuscript from the UCI Machine Learning Repository. The dataset contains 400 observations, 24 predictor variables, and one class variable indicating CKD or non-CKD. The proposed framework was defined by combining prior odds with sequential likelihood evidence and by using explicit stopping boundaries for diagnostic decisions. Performance was assessed using accuracy, sensitivity, specificity, precision, F1-score, and area under the receiver operating characteristic curve (AUC). Based on the reported results, the proposed PS-BMSPRT achieved an accuracy of 95.1%, sensitivity of 94.3%, specificity of 95.8%, F1-score of 94.7%, and AUC of 0.98. The machine-learning comparison reported XGBoost and Random Forest as the strongest classifiers. However, the exact hyperparameters, training/validation split, and prior hyperparameters were not reported in the submitted manuscript and must be verified against the original analysis before publication.

Keywords: Bayesian MSPRT; informative priors; patient-specific risk; chronic kidney disease; sequential probability ratio test; machine learning; ROC-AUC.

1. Introduction

Kidney disease remains a major global public health challenge, affecting millions of individuals worldwide and contributing substantially to morbidity, mortality, and healthcare expenditure. Chronic Kidney Disease (CKD) and Acute Kidney Injury (AKI) are particularly concerning due to

their progressive nature and often asymptomatic early stages, which frequently result in delayed diagnosis and treatment (Hastie *et al,* 2009). According to the Global Burden of Disease study, kidney diseases have experienced a significant increase in prevalence and mortality over recent decades, emphasising the need for more accurate and timely diagnostic strategies (GBD Chronic Kidney Disease Collaboration, 2020; Iftikhar *et al,* 2023). Early detection is critical because prompt intervention can slow disease progression, reduce complications, and improve patient outcomes (Levey & Coresh, 2012; Kourou *et al* 2015; Jain & Singh, 2018).

Traditional diagnostic methods for kidney disease rely on biomarkers such as serum creatinine, estimated glomerular filtration rate (eGFR), albuminuria, and imaging assessments (Nguyen *et al,* 2025; Pedregosal *et al* 20110. While these approaches have demonstrated clinical utility, they often employ fixed decision thresholds that do not adequately account for patient heterogeneity, disease progression patterns, or evolving clinical evidence (Levey *et al* 2003; Stevens & Levin, 2013; Goldstein *et al,* 2017). Consequently, diagnostic decisions may be delayed or associated with increased rates of false-positive and false-negative classifications, potentially compromising treatment effectiveness and resource utilization Breiman (2001).

Sequential statistical methodologies offer an attractive alternative by enabling decisions to be made as data accumulate rather than after a predetermined sample size has been reached. The Sequential Probability Ratio Test (SPRT), introduced by Wald 1947, Witten *et al. (*2016), and Zhang *et al, (*2019), is recognised for its efficiency in hypothesis testing, minimising the expected number of observations while maintaining specified error rates (Stevens & Levin, 2013). However, conventional SPRT frameworks are typically based on fixed prior assumptions and population-level parameters, limiting their adaptability to individual patient characteristics and dynamic clinical environments.

Bayesian approaches provide a principled framework for incorporating prior knowledge and continuously updating diagnostic probabilities as new patient information becomes available (Chem & Guestrain, 2016). Integrating patient-specific clinical histories, demographic characteristics, laboratory measurements, and risk factors, Bayesian methods can generate personalised posterior probabilities that better reflect an individual's disease status (Spiegelhalter, Abrams, & Myles, 2004; Guyon & Elisseeff, 2003).

Recent advances in personalised medicine have further highlighted the importance of individualised decision-support systems capable of adapting to patient-specific variability and uncertainty (Jameson & Longo, 2015; Cortes & Vapnik, 1995). These considerations, this study proposes an Enhanced Patient-Specific Bayesian Modified Sequential Probability Ratio Test (PS-BMSPRT) for personalised kidney disease diagnosis (Chandrashekar & Safin, 2014).

The proposed framework extends the traditional SPRT by incorporating Bayesian updating mechanisms and patient-specific prior distributions, allowing diagnostic decisions to adapt dynamically to each patient's evolving clinical profile (Smyth, 2001; Dua & Graft, 2019). By leveraging individualised risk information and sequential evidence accumulation, the proposed method aims to improve diagnostic accuracy, reduce decision latency, and enhance clinical decision-making compared with conventional approaches (Aljaaf *et al* 2018; Bansal & Kats, 2019).

The significance of this research lies in its potential to bridge the gap between statistical decision theory and personalised healthcare. Through the integration of Bayesian inference and modified sequential testing procedures, the proposed approach seeks to provide a robust and efficient diagnostic framework capable of supporting early and reliable kidney disease detection (Charma *et al.,* 2022). Ultimately, the study contributes to the growing body of research on precision medicine

by demonstrating how patient-specific probabilistic models can be utilised to optimise diagnostic performance and improve healthcare outcomes (Rajaraman & Ullman, 2021).

2. Materials and Methods

The analysis used the Chronic Kidney Disease dataset from the UCI Machine Learning Repository, which is the dataset source cited in the submitted manuscript (Dua & Graff, 2019). The UCI record describes 400 instances, 24 predictor variables, and one class variable, with mixed numerical and categorical attributes and missing values. The target variable has two classes: CKD and not-CKD. The dataset is publicly available and was originally collected over approximately two months.

It is directly proportional to the root mean square of the observations and is used in constructing likelihood ratios for the MSPRT procedure.

Slope will be given as:

$$S = \log \frac{\theta_1^3}{\theta_2^3} \left(\frac{1}{\theta_1^3} - \frac{1}{\theta_2^3} \right)^{-1} \quad (1).$$

And the intercept y_1 is equal to

$$h_1 = \frac{\beta}{1-\alpha} \left(\frac{1}{\theta_1^3} - \frac{1}{\theta_2^3} \right)^{-1} \quad (2).$$

And the intercept y_2 is equal to

$$h_2 = \frac{1-\beta}{\alpha} \left(\frac{1}{\theta_1^3} - \frac{1}{\theta_2^3} \right)^{-1} \quad (3)$$

The number of observations is measured along the horizontal axis since both h_1 and h_2 Linear functions of n will lie on a straight line. y_1 and y_2 The two lines are parallel to each other; this estimator is used in constructing the likelihood ratio for the MSPRT procedure.

Let α be the maximum type-I error probability and β be the maximum type-II error probability. The conventional SPRT likelihood-ratio boundaries are

$$A=(1-\beta)/\alpha,\ B = \beta/(1-\alpha) \quad (4)$$

The models were evaluated using accuracy, sensitivity (recall), specificity, precision, F1-score, and AUC. For a binary classifier, let TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively.

$$\text{Sensitivity} = TP/(TP + FN) \quad (5)$$

$$\text{Specificity} = TN/(TN + FP) \quad (6)$$

$$\text{Precision} = TP/(TP + FP) \quad (7)$$

$$\text{F1-score} = 2 \times (\text{Precision} \times \text{Sensitivity})/(\text{Precision} + \text{Sensitivity}) \quad (8)$$

$$\text{Accuracy} = (TP+ TN)/(TP + TN + FP + FN) \quad (9)$$

The false-positive rate used in ROC analysis is FPR = 1 − Specificity. The AUC is the area under the ROC curve and summarises discrimination across classification thresholds. The confidence interval for AUC should be reported using the method actually used in the analysis (e.g., DeLong or bootstrap); the original manuscript did not identify the method.

3. Results and Discussions

Random Forest and XGBoost achieved the highest performance. SVM also performed strongly. KNN showed limitations due to high dimensionality. Ensemble methods outperform single classifiers due to better handling of high-dimensional data.

Table 1. PS-BMSPRT

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	AUC
Traditional SPRT	84.5	82.3	86.1	83.4	0.87
Bayesian SPRT	89.2	87.8	90.5	88.6	0.92
Proposed PS-BMSPRT	95.1	94.3	95.8	94.7	0.98

Table 1. Diagnostic performance comparison of the SPRT models.

Table 2. SVM and Naïve Bayes

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	95%	94%	93%	93%
SVM	98%	97%	96%	96%
KNN	93%	92%	91%	91%
Naïve Bayes	97%	96%	95%	95%
Random Forest	99–100%	99%	99%	99%
XGBoost	99–100%	99%	99%	99%

Table 2. Reported machine-learning performance comparison.

Reproduces the machine-learning results reported in the submitted manuscript. XGBoost and Random Forest were reported as the strongest models, followed by SVM and Naïve Bayes. KNN and Logistic Regression showed comparatively lower values. These results can support the selection of high-performing machine-learning components only if the implementation and validation procedure are documented and reproducible.

Table 3. ROC/AUC Results

Model	AUC
Traditional SPRT	0.87
Bayesian SPRT	0.92
Proposed PS-BMSPRT	0.98

Table 3. AUC comparison of the SPRT models.

The reported AUC of 0.98 for PS-BMSPRT indicates excellent discrimination between CKD and non-CKD cases. The AUC values of 0.87 and 0.92 for traditional and Bayesian SPRT, respectively, indicate progressively improved discrimination. The manuscript should include the actual ROC curves generated from the validated analysis and should state how confidence intervals were obtained.

4. Conclusion

The study concluded that the enhanced PS-BMSPRT model provided a statistically robust and clinically promising approach for personalised kidney disease diagnosis. Its superior performance suggests that combining Bayesian inference with sequential testing and patient-specific information can improve diagnostic efficiency and reliability. The proposed model therefore represents a valuable contribution to the development of intelligent and personalised diagnostic systems for kidney disease. Future studies should validate the model using larger, diverse, and real-time clinical datasets and investigate its integration with machine learning and clinical decision-support systems to further improve its applicability in routine healthcare practice. The findings suggest that integrating machine learning into clinical decision systems can significantly enhance early detection and patient outcomes.

References

Aljaaf, A. J., Al-Jumeily, D., Hussain, A., Dawson, T., & Fergus, P. (2018). Predicting the likelihood of chronic kidney disease using machine learning techniques. *Healthcare Informatics Research, 24*(3), 185–194. https://doi.org/10.4258/hir.2018.24.3.185

Bansal, N., & Katz, R. (2019). Machine learning approaches for predicting chronic kidney disease progression. *Kidney International Reports, 4*(3), 321–329. https://doi.org/10.1016/j.ekir.2018.11.005

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering, 40*(1), 16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*(3), 273–297. https://doi.org/10.1007/BF00994018

Dua, D., & Graff, C. (2019). UCI machine learning repository. *University of California, Irvine*. https://archive.ics.uci.edu

GBD Chronic Kidney Disease Collaboration. (2020). Global, regional, and national burden of chronic kidney disease, 1990–2017: A systematic analysis for the Global Burden of

Goldstein, B. A., Navar, A. M., Pencina, M. J., & Ioannidis, J. P. A. (2017). Opportunities and challenges in developing risk prediction models with electronic health records data. *Circulation, 135*(5), 453–467. https://doi.org/10.1161/CIRCULATIONAHA.116.023663

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research, 3*, 1157–1182.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning* (2nd ed.). Springer.

Iftikhar, H., Fatima, A., & Qureshi, S. (2023). Comparative analysis of machine learning algorithms for chronic kidney disease prediction. *Sustainability, 15*(3), 2754. https://doi.org/10.3390/su15032754

Jain, D., & Singh, V. (2018). Feature selection and classification systems for chronic disease prediction: A review. *Egyptian Informatics Journal, 19*(3), 179–189. https://doi.org/10.1016/j.eij.2018.03.002

Jameson, J. L., & Longo, D. L. (2015). Precision medicine—Personalised, problematic, and promising. *New England Journal of Medicine, 372*(23), 2229–2234. https://doi.org/10.1056/NEJMsb1503104

Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal, 13*, 8–17. https://doi.org/10.1016/j.csbj.2014.11.005

Levey, A. S., & Coresh, J. (2012). Chronic kidney disease. *The Lancet, 379*(9811), 165–180. https://doi.org/10.1016/S0140-6736(11)60178-5

Levey, A. S., Coresh, J., Baulk, E., Kausz, A. T., Levin, A., & Steffes, M. W. (2003). National Kidney Foundation practice guidelines for chronic kidney disease. *Annals of Internal Medicine, 139*(2), 137–147. https://doi.org/10.7326/0003-4819-139-2-200307150-00013

Nguyen, D. P., Tran, T. H., & Le, Q. H. (2025). Machine learning techniques for chronic kidney disease diagnosis: A comparative study. *BMC Medical Informatics and Decision Making*. https://pubmed.ncbi.nlm.nih.gov/40735336/

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., & Grisel, O. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.

Rajaraman, A., & Ullman, J. D. (2011). *Mining of massive datasets*. Cambridge University Press.

Sharma, S., Singh, A., & Verma, P. (2022). Machine learning approaches for chronic kidney disease diagnosis: A comprehensive review. *Smart Health, 26*, 100321. https://doi.org/10.1016/j.smhl.2022.100321

Smyth, P. (2001). Data mining: Data analysis on a grand scale? *Statistical Methods in Medical Research, 10*(5), 309–325. https://doi.org/10.1191/096228001680193878

Spiegelhalter, D. J., Abrams, K. R., & Myles, J. P. (2004). *Bayesian approaches to clinical trials and health-care evaluation*. John Wiley & Sons.

Stevens, P. E., & Levin, A. (2013). Evaluation and management of chronic kidney disease: Synopsis of the Kidney Disease: Improving Global Outcomes 2012 Clinical Practice Guideline. *Annals of Internal Medicine, 158*(11), 825–830. https://doi.org/10.7326/0003-4819-158-11-201306040-00007

Wald, A. (1947). *Sequential analysis*. John Wiley & Sons.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data mining: Practical machine learning tools and techniques* (4th ed.). Morgan Kaufmann.

Zhang, Z., Ho, K. M., & Hong, Y. (2019). Machine learning for the prediction of chronic kidney disease progression. *Nephrology, 24*(2), 131–137. https://doi.org/10.1111/nep.13435