FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 10 No. 14, July, 2026, pp 152 – 156
DOI: https://doi.org/10.33003/fjs-2026-1014-5701

A Systematic Review of Chronic Kidney Disease Prediction in Nigeria using Machine Learning Techniques

***[1]Hauwa Ahmad Amshi, [2]Saratu Yusuf Ilu and [3]Usman Muhammad Kadai**

[1]Federal University Gashua, Yobe State, Nigeria.
[2]Bayero University Kano, Kano State, Nigeria.
[3]IDHL Group, Nigeria.

*Corresponding authors' email: hauwa.amshi@fugashua.edu.ng

Received: 07 Aug. 2026
Accepted: 08 Aug. 2026
Published: 17 Aug. 2026

ABSTRACT
Chronic kidney disease (CKD) is a major public health problem in Nigeria, where delayed presentation, limited diagnostic capacity and low awareness contribute to under-diagnosis. Early identification of individuals at risk may help delay progression to end-stage renal disease. Machine learning (ML) provides an opportunity to improve risk stratification using routinely collected clinical data. This systematic review synthesized studies that used Nigerian clinical data to develop supervised ML models for CKD prediction or diagnosis. A systematic search of PubMed, Scopus, IEEE Xplore, Google Scholar and African Journals Online was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines for studies published between January 2010 and April 2025. Methodological quality and risk of bias were assessed using the Prediction Model Risk of Bias Assessment Tool (PROBAST). Six empirical studies met the eligibility criteria. Reported model accuracy ranged from 81.3% to 99.4%, with Random Forest, Gradient Boosting, Deep Neural Networks and hybrid approaches among the methods used. Serum creatinine, age, blood pressure, blood urea nitrogen and haemoglobin were frequently reported predictors. Major limitations included small and heterogeneous datasets, limited interpretability, inadequate external validation and limited translation of models into routine clinical practice. Future research should prioritize multicentre prospective datasets, transparent reporting, external validation and explainable ML approaches to support clinically useful and equitable CKD prediction in Nigeria.

Keywords: Chronic kidney disease; machine learning; Nigeria; prediction models; early diagnosis; artificial intelligence

INTRODUCTION

Chronic kidney disease (CKD) is a significant global public health problem, affecting approximately 10–15% of adults worldwide and contributing substantially to morbidity, mortality and economic burden (Bikbov *et al.*, 2020). In Nigeria, CKD prevalence is increasing in the context of an epidemio-syndemic involving rising non-communicable diseases, late presentation, suboptimal access to care and limited diagnostic capacity (Apiyanteide *et al.*, 2023; Olanrewaju *et al.*, 2020). Early diagnosis remains important for slowing disease progression and reducing complications, particularly in resource-constrained settings. CKD progresses through five stages, with stage 5 representing end-stage renal disease requiring renal replacement therapy. Because early CKD may be asymptomatic, timely identification remains challenging (Iftikhar *et al.*, 2025).

Conventional assessment relies on measures such as estimated glomerular filtration rate (eGFR) and proteinuria, although these measures may not capture all early changes in kidney function. Recent advances in machine learning (ML) and artificial intelligence (AI) offer additional approaches for identifying complex patterns in clinical data and stratifying individuals according to CKD risk (Khalid *et al.*, 2024; Mbunge & Batani, 2023). In CKD prediction, ML models can analyse combinations of demographic, clinical and laboratory variables to support risk classification and potentially improve clinical decision-making (Zheng *et al.*, 2024).

In Nigeria, the clinical application of ML for CKD prediction remains limited, although available studies have reported promising results using variables such as blood pressure, age, serum creatinine, diabetes and hypertension. For example, Ibrahim Iliyas *et al.* (2021) reported a Deep Neural Network model for CKD prediction with an accuracy of 98%. Nevertheless, implementation in routine care is constrained by limited collaboration between clinicians and data scientists, insufficient ML literacy among healthcare workers, data privacy concerns and the risk of algorithmic bias (Owoyemi *et al.*, 2020). Socioeconomic and cultural factors, including health literacy, trust in digital technologies and healthcare infrastructure, must also be considered to ensure that predictive tools are effective and equitable.

This systematic review therefore critically synthesizes available studies on ML-based CKD prediction using Nigerian data. It examines the algorithms used, predictive performance, methodological quality, feature-selection strategies and commonly reported predictors, while identifying research gaps that need to be addressed before these models can be translated into routine clinical practice.

MATERIALS AND METHODS

Study Design

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines were followed to improve transparency and reporting of the review (Page *et al.*, 2021). Because the included studies differed substantially in

their clinical data sources, algorithmic architectures, software and feature-selection methods, a qualitative narrative synthesis was used. The heterogeneity of the available evidence and the limited reporting of confidence intervals or standard errors precluded a quantitative meta-analysis.

Search Strategy

A systematic search of PubMed, Scopus, IEEE Xplore, Google Scholar and African Journals Online (AJOL) was conducted. The publication period was January 2010 to April 2025. Search strings were adapted to the syntax of each database and combined the following concepts: ("chronic kidney disease" OR "CKD") AND ("machine learning" OR "ML" OR "artificial intelligence" OR "AI") AND ("prediction" OR "diagnosis") AND ("Nigeria"). Forward and backward citation searching of eligible studies was also conducted in Google Scholar. Retrieved references were managed in EndNote, with automated and manual deduplication.

Eligibility Criteria

Studies were eligible if they were empirical investigations that used ML or deep learning (DL) approaches to predict or diagnose CKD or related renal dysfunction; used data generated in clinical settings or from community-dwelling populations in Nigeria; reported classification-performance measures such as accuracy, sensitivity, specificity or area under the receiver operating characteristic curve (AUC-ROC); were published in English; and fell within the January 2010 to April 2025 period. Conceptual papers, meta-analyses, editorials, commentaries, letters and conference abstracts without empirical validation data were excluded. Studies relying solely on external datasets, including non-Nigerian repositories such as the UCI machine-learning repository, were excluded, as were models intended only for non-diagnostic applications such as treatment optimization or drug discovery and animal or in-vitro studies.

Selection Process

After automated deduplication, two reviewers independently screened titles and abstracts to identify potentially eligible studies. Full texts of potentially eligible articles were then retrieved and assessed against the inclusion criteria. Disagreements were resolved through discussion and, where necessary, consultation with a third reviewer. No eligible study was lost at the screening or full-text stages.

Data Extraction and Quality Assessment

Data were extracted independently by two reviewers using a standardized Excel template. Extracted variables included publication year, geographic setting, sample size, data-source type, algorithmic framework, data-selection methods, validation methods such as cross-validation or hold-out validation, and reported performance measures. Methodological quality and risk of bias were assessed using the Prediction Model Risk of Bias Assessment Tool (PROBAST) (Wolff *et al.*, 2019), considering the Participants, Predictors, Outcome and Analysis domains. Each domain was classified according to the reported level of risk, and an overall assessment was made for each study.

RESULTS AND DISCUSSION

Study Selection Flow

A total of 321 records were identified from the electronic databases. After removal of 89 duplicates, 232 titles and abstracts were screened. Of these, 195 were excluded because they did not use ML or AI (n = 52), did not focus on CKD (n = 48), used non-Nigerian data (n = 45), were review or editorial papers (n = 30), were published before 2010 (n = 12), or were not in English (n = 8). Thirty-seven full-text reports were assessed for eligibility. A further 31 studies were excluded because they did not report classification performance or relied on standardized external datasets such as the UCI repository. Six empirical studies met all eligibility criteria and were included in the qualitative synthesis.

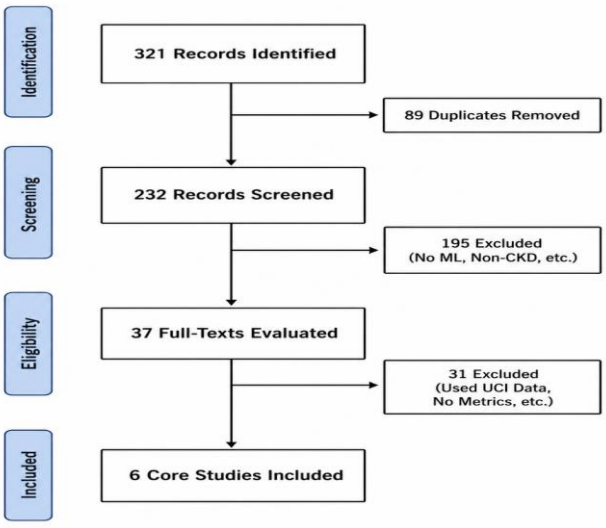

Figure 1: PRISMA diagram showing the selection process of included studies.
PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Methodological Quality and Risk of Bias (PROBAST)

The risk-of-bias assessment of the six included studies was summarized across the four PROBAST domains. Osofisan *et al*. (2011) were assessed as having moderate risk in the Participants, Predictors and Outcome domains and low risk in the Analysis domain. Ibrahim Iliyas *et al.* (2021) had low risk

in the Participants and Predictors domains, moderate risk in Analysis and high risk in Outcome. Adunni *et al*. (2022) had low risk across all four domains. Aliyu *et al*. (2025) had low risk in Participants, Predictors and Analysis and moderate risk in Outcome. Bagbe *et al*. (2025) had moderate risk in Predictor selection and Analysis and low risk in Participants and Outcome. Across the studies, recurrent concerns included limited reporting of model calibration, insufficient description of clinical data handling and reliance on internal rather than external validation. More recent publications showed greater attention to methodological and reporting standards.

Study	Participants	Predictors	Outcome	Analysis	Overall Quality
Osofisan et al. (2011)	Moderate	Moderate	Moderate	Low	Moderate Risk
Iliyas (2020)	Low	Low	Moderate	Low	Low-to-Moderate
Iliyas, Rambo, et al. (2021)	Low	Low	Moderate	Low	Low-to-Moderate
Adunni et al. (2022)	Low	Low	Low	Low	Low Risk
Aliyu et al. (2025)	Low	Low	Moderate	Low	Low-to-Moderate
Bagbe et al. (2025)	Moderate	Low	Moderate	Low	Moderate Risk

Figure 2: Visual summary of the methodological quality assessment conducted using the PROBAST framework

Algorithmic Implementations and Predictive Performance

The algorithms reported in the included studies can be grouped into conventional ML, deep learning and ensemble-learning approaches. Conventional ML methods included Decision Trees, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Naïve Bayes and Logistic Regression, with feature engineering used to improve classification. Deep-learning approaches included Multi-Layer Perceptrons (MLP) and Deep Neural Networks (DNN), which were used to model nonlinear relationships among clinical variables. Ensemble approaches included Random Forest, Gradient Boosting, AdaBoost and stacking methods. The reported performance of the models varied according to the algorithm, data characteristics and preprocessing procedures.

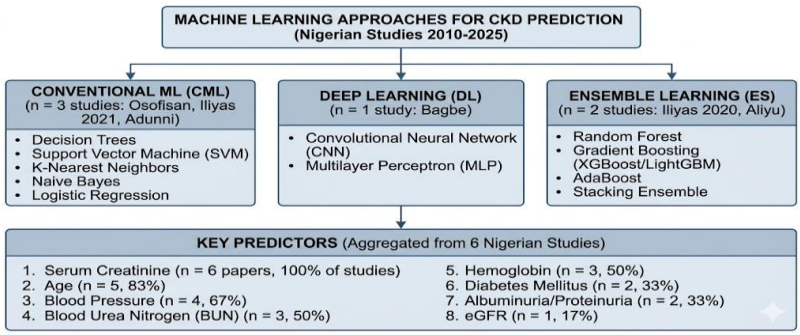

Figure 3: Machine learning approaches used for CKD prediction.

Table 1 Summarizes the Principal Algorithms, Clinical Settings and Peak Reported Accuracy of the Included Studies

Author (Year)	Core Algorithmic Framework	Data Source / Clinical Setting	Peak Reported Accuracy
Osofisan et al. (2011)	Artificial Neural Network (ANN)	University College Hospital (UCH), Ibadan	Not fully reported
Iliyas et al. (2021a)	Deep Neural Network (DNN)	Bade General Hospital, Yobe State	98.0%
Iliyas et al. (2021b)	DNN, ANN, KNN, Naïve Bayes and Logistic Regression	Gashua General Hospital, Yobe State	98.0%
Adunni et al. (2022)	Ant Colony Optimization–KNN (ACO-KNN)	Multicentre (three hospitals, South-West Nigeria)	99.4%
Aliyu et al. (2025)	Random Forest / ensemble	Sokoto EHR data	81.3%
Bagbe et al. (2025)	Random Forest / deep learning	Ondo State	98.5%

The six included studies reported peak accuracies ranging from 81.3% to 99.4%. The highest reported accuracy was obtained by the ACO-KNN hybrid framework (99.4%), while Random Forest and deep-learning approaches also demonstrated high reported accuracy. These results should be interpreted cautiously because the studies differed in sample size, validation strategy, data source and reporting completeness.

Feature Engineering and Frequency of Key Predictors

Feature selection and dimensionality-reduction methods were used to reduce noise, computational burden and potential overfitting. Reported approaches included Principal Component Analysis (PCA), Recursive Feature Elimination (RFE) and metaheuristic optimization such as Ant Colony Optimization. Across the six included studies, serum creatinine was reported in all six studies (100%), age in five

(83%), blood pressure in four (67%), blood urea nitrogen (BUN) in three (50%), haemoglobin in three (50%), diabetes mellitus in two (33%), albuminuria or proteinuria in two (33%), and eGFR in one study (17%).

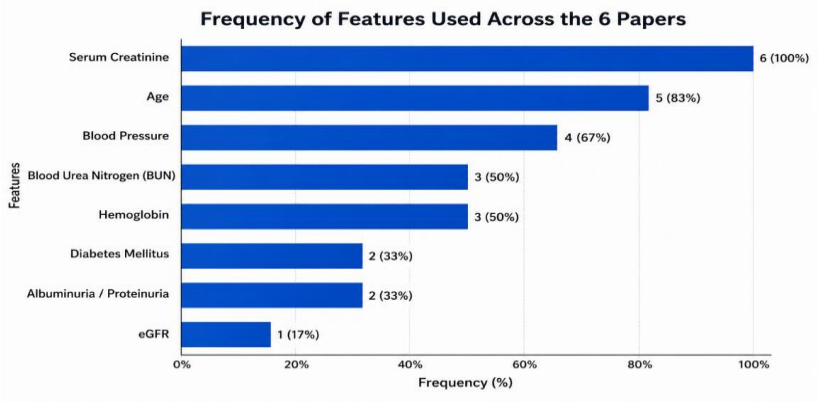

Figure 4: Features used across selected papers

Only one of the six included studies (17%) reported eGFR, despite its importance in CKD staging. The apparent reliance on raw serum creatinine may reflect differences in laboratory standardization, availability of appropriate eGFR equations and resource constraints. Some studies also incorporated additional biochemical parameters, including serum bicarbonate, into their predictive models.

Discussion

The reviewed studies indicate that ML workflows developed using Nigerian clinical datasets can achieve high reported classification performance for CKD prediction. The findings also emphasize the value of locally generated data because disease presentation, comorbidities and sociodemographic characteristics may differ across populations. However, the high reported performance should not be interpreted as evidence of clinical readiness because most studies were based on relatively small datasets and lacked robust external validation.

Geographic and socioeconomic data skew remain important limitations. Available datasets are frequently derived from tertiary or teaching hospitals located in urban areas, including University College Hospital, Ibadan and other major referral centres. Consequently, rural and peri-urban populations may be underrepresented. This limitation is particularly important in Nigeria, where access to laboratory testing and other diagnostic services varies substantially across settings.

Many of the reviewed datasets were cross-sectional or based on short retrospective periods. Such datasets can support binary classification of CKD versus non-CKD, but they provide limited information about the longitudinal progression of CKD stages. Prospective longitudinal datasets would therefore be valuable for developing models that predict disease progression rather than only current disease status.

Interpretability is another important barrier. Deep-learning and ensemble models may achieve strong predictive performance but can be difficult to interpret because of their nonlinear decision processes. Limited transparency may reduce clinicians' confidence in model recommendations. Explainable artificial intelligence approaches could help address this barrier by providing clinically meaningful explanations for predictions.

The reviewed models also focused predominantly on clinical and laboratory variables and gave limited attention to broader socioeconomic and environmental determinants that may influence the equity and practical utility of CKD screening. Future studies should therefore incorporate relevant contextual factors where appropriate while maintaining data quality, privacy and clinical validity.

To reduce the research-to-implementation gap, future work should prioritize multicentre prospective datasets, transparent reporting, appropriate handling of class imbalance, external validation across geographic settings and explainable ML interfaces. Such measures would improve the generalizability, trustworthiness and potential clinical utility of CKD prediction models in Nigeria.

CONCLUSION

Supervised machine-learning algorithms show substantial potential for supporting early CKD screening and improving the allocation of healthcare resources in Nigeria. However, the evidence base remains limited by small and heterogeneous datasets, restricted geographic representation, incomplete reporting, limited external validation and concerns about model interpretability. Future research should focus on multicentre prospective datasets, standardized data collection, external validation across diverse Nigerian settings and clinically interpretable ML approaches. Collaboration among clinicians, data scientists, health-system stakeholders and policymakers will be essential for translating promising predictive models into safe, equitable and sustainable clinical tools.

REFERENCES

Adunni, O. O., Olusayo, O. E., Olatunde, O. S., Segun, S. O., & Aina, S. (2022). ACO-KNN predictive model for diagnosis of chronic kidney disease. International Journal of Research and Innovation in Applied Science.

Aliyu, A. M., Mumin, K. A., Abubakar, A. B., Abba, A. S., & Ibrahim, A. (2025). Performance evaluation of machine learning algorithm using chronic kidney disease (CKD) dataset. FUDMA Journal of Sciences, 9(12), 72–80. https://doi.org/10.33003/fjs-2025-0912-4025

Apiyanteide, F., Eu, N., Cc, O., & Digban, K. (2023). Systematic review and meta-analysis of the epidemiology of chronic kidney disease in Nigeria, 11(12), 690–709.

Bagbe, A., Akomolafe, A. A., & Bagbe, A. S. (2025). Development and evaluation of a machine learning model for

predicting terminal diseases. Asian Journal of Probability and Statistics, 27(4), 157–165. https://doi.org/10.9734/ajpas/2025/v27i4747

Bakare, O., Agaba, E., Gimba, Z., Abene, E., Imoh, L., & Maji, J. (2025). Assessing renal function in chronic kidney disease: A comparative evaluation of glomerular filtration rate prediction equations in the North-Central Region of Nigeria. Cureus, 17. https://doi.org/10.7759/cureus.84577

Barrera-García, J., Cisternas-Caneo, F., Crawford, B., Gómez Sánchez, M., & Soto, R. (2024). Feature selection problem and metaheuristics: A systematic literature review about its formulation, evaluation and applications. Biomimetics, 9(1). https://doi.org/10.3390/biomimetics9010009

Bikbov, B., Purcell, C. A., Levey, A. S., *et al.* (2020). Global, regional, and national burden of chronic kidney disease, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. The Lancet, 395(10225), 709–733. https://doi.org/10.1016/S0140-6736(20)30045-3

Chawla, N. V., Bowyer, K. W., & Hall, L. O. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321–357.

Delrue, C., De Bruyne, S., & Speeckaert, M. M. (2024). Application of machine learning in chronic kidney disease: Current status and future prospects. Biomedicines, 12(3). https://doi.org/10.3390/biomedicines12030568

Ibrahim Iliyas, I., Rambo, S. I., Dauda, A. B., & Tasiu, S. (2021a). Prediction of chronic kidney disease using deep neural network. FUDMA Journal of Sciences, 4(4), 34–41. https://doi.org/10.33003/fjs-2020-0404-309

Iftikhar, H., Hashem, A. F., Qureshi, M., & Rodrigues, P. C. (2025). Clinical application of machine learning models for early-stage chronic kidney disease detection. Diagnostics, 15(20). https://doi.org/10.3390/diagnostics15202610

Iliyas, I. I., Boukari, S., & Ya, A. (2025). Recent trends in prediction of chronic kidney disease using different learning approaches: A systematic literature review. Journal of Medical Artificial Intelligence. https://doi.org/10.21037/jmai-24-256

Iliyas, I. I., Isah, R. S., Ali, B. D., & Andra, U. (2021b). Comparative analysis of some prominent machine learning algorithms for the prediction of chronic kidney disease. Academy Journal of Science and Engineering, 15(1).

Islam, M. A., Majumder, M. Z. H., & Hussein, M. A. (2023). Chronic kidney disease prediction based on machine learning algorithms. Journal of Pathology Informatics, 14, 100189. https://doi.org/10.1016/j.jpi.2023.100189

Khalid, F., Alsadoun, L., Khilji, F., *et al.* (2024). Predicting the progression of chronic kidney disease: A systematic review of artificial intelligence and machine learning approaches. Cureus, 16(5), e60145. https://doi.org/10.7759/cureus.60145

Mbunge, E., & Batani, J. (2023). Application of deep learning and machine learning models to improve healthcare in sub-Saharan Africa: Emerging opportunities, trends and implications. Telematics and Informatics Reports, 11, 100097. https://doi.org/10.1016/j.teler.2023.100097

Olanrewaju, T. O., Aderibigbe, A., Popoola, A. A., *et al.* (2020). Prevalence of chronic kidney disease and risk factors in North-Central Nigeria: A population-based survey. BMC Nephrology, 21(1), 467. https://doi.org/10.1186/s12882-020-02126-8

Osofisan, A. O., Adeyemo, O. O., Sawyerr, B. A., & Eweje, O. (2011). Prediction of kidney failure using artificial neural networks. European Journal of Scientific Research, 61(4), 487–492.

Owoyemi, A., Owoyemi, J., Osiyemi, A., & Boyd, A. (2020). Artificial intelligence for healthcare in Africa. Frontiers in Digital Health, 2, 6. https://doi.org/10.3389/fdgth.2020.00006

Page, M. J., McKenzie, J. E., Bossuyt, P. M., *et al.* (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71

Rashed-Al-Mahfuz, M., Haque, A., Azad, A., *et al.* (2021). Clinically applicable machine learning approaches to identify attributes of chronic kidney disease (CKD) for use in low-cost diagnostic screening. IEEE Journal of Translational Engineering in Health and Medicine, 9, 4900511. https://doi.org/10.1109/JTEHM.2021.3073629

Wolff, R. F., Moons, K. G. M., Riley, R. D., *et al.* (2019). PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. Annals of Internal Medicine, 170(1), 51–58. https://doi.org/10.7326/M18-1376

Zheng, J.-X., Li, X., Zhu, J., *et al.* (2024). Interpretable machine learning for predicting chronic kidney disease progression risk. DIGITAL HEALTH, 10, 20552076231224224. https://doi.org/10.1177/20552076231224225