International Journal of Basic and Applied Science 12 (1) (2023) 33-38

Published by: IOCSCIENCE

International Journal of Basic and Applied Science

Journal homepage: www.ijobas.pelnus.ac.id

Exploring the role of urine analysis in early detection of chronic kidney disease

Frank H
NPA, Rockville, USA

Article Info

Article history:

Received May 20, 2023
Revised Jun 16, 2023
Accepted Jun 30, 2023

Keywords:

Chronic Kidney Disease;
Clinical Significance;
Diagnostic Model;
Early Detection;
Performance Evaluation.

ABSTRACT

This study focuses on the development and validation of a urine-based diagnostic model for early detection of Chronic Kidney Disease (CKD). Only urine indicators, hypertension, diabetes mellitus, coronary artery disease, appetite, pedal edema, and anemia, are selected based on their relevance to CKD. A decision tree algorithm is utilized for model development, with specific parameters set for optimal performance. The model is trained and evaluated using two datasets, demonstrating promising results in terms of 100% true positive and true negative rates in validation study. The findings highlight the potential clinical significance and applicability of the developed model for timely interventions in CKD patients.

Corresponding Author:

Frank (D-C) H,
NPA,
N.P.A. Rockville,
15800 Crabbs Branch Way, MD, USA
Email: hillsbros.formosa@gmail.com

1. INTRODUCTION

Chronic Kidney Disease (CKD) is a progressive condition characterized by a gradual decline in kidney function over time, leading to a range of complications and increased morbidity and mortality rates [1]. This disease has become a significant public health concern globally, with a rising incidence and prevalence in both developed and developing countries [2]. Early detection and intervention play a crucial role in managing CKD effectively and preventing its progression to end-stage renal disease (ESRD). Timely diagnosis allows for the implementation of interventions such as lifestyle modifications, pharmacotherapy, and appropriate referral to nephrology specialists, ultimately improving patient outcomes [3]. However, CKD is often asymptomatic in its early stages, making early detection challenging. Therefore, the development of accurate and non-nvasive diagnostic tools is of paramount importance in improving clinical outcomes for CKD patients. Serum Creatinine Level. "Serum creatinine level is a widely used indicator for assessing renal function in CKD diagnosis. Elevated levels of serum creatinine indicate impaired kidney function and decreased glomerular filtration rate (GFR)". Calculation Methods for Glomerular Filtration Rate (GFR) such as MDRD and CKD-EPI Equations. "GFR estimation equations, such as the Modification of Diet in Renal Disease (MDRD) and Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equations, are commonly utilized to estimate GFR based on serum creatinine levels, age, gender, and race". Renal Histopathology Examination. "Renal histopathology examination involves the microscopic evaluation of kidney tissue samples obtained through biopsy. It provides crucial information about the structural abnormalities in glomeruli, tubules, and interstitium, aiding in the confirmation and classification of renal damage".

Imaging Techniques (Ultrasound, CT Scans, MRI, etc.)."Imaging techniques such as ultrasound, computed tomography (CT) scans, and magnetic resonance imaging (MRI) play a significant role in identifying structural abnormalities and defects in the kidneys. These non-invasive imaging modalities assist in the diagnosis and evaluation of renal pathologies" [4]. Urine Examination Indicators.

Quantitative Assessment of Urinary Protein."Quantitative measurement of urinary protein, typically expressed as the protein-to-creatinine ratio (PCR), is an essential indicator for evaluating renal damage. Increased levels of urinary protein (proteinuria) are suggestive of glomerular dysfunction and are considered a hallmark of CKD" [5].

Urine Specific Gravity. "Urinary specific gravity (SG) reflects the concentration of solutes in the urine and can provide insights into the kidney's ability to concentrate urine. Abnormal SG levels may indicate impaired renal function" [6].

Presence of Red Blood Cells (RBCs) and White Blood Cells (WBCs) in Urine. "The presence of red blood cells (hematuria) and white blood cells (pyuria) in urine can signify underlying kidney damage or inflammation. Urine microscopy is commonly employed to detect the presence of these cells"[7].

Table 1. Summary of Key Research Topics and Methods in Chronic Kidney Disease (CKD) Diagnosis

Research Topic	Key Points	Research Methods
CKD Diagnosis Criteria and Background	Global prevalence and background of CKD [3]	Literature review, global prevalence analysis
Assessment of Renal Dysfunction	Serum creatinine and GFR calculation methods [1][8][9]	KDIGO guidelines, CKD-EPI and MDRD formulas, literature review
Application of Urine Markers in CKD Diagnosis	Significance of urine protein, specific gravity, and RBC and WBC counts [10][11][12]	Community-based urine testing, urine microscopy, literature review
Impact of WBC and RBC in Urine on CKD Diagnosis	Importance of WBC and RBC counts in renal disease assessment [13][14][15][16]	Urine microscopy examination, literature review
Urinalysis Application in CKD Diagnosis	Detecting proteinuria and other indicators in urine [17][18]	Urine analysis indicators and application, literature review
Hematuria and CKD Diagnosis	Correlation betweenhematuria and CKD diagnosis [19][20]	Urine analysis indicators and application, literature review

2. RESEARCH METHOD

2.1. Selection of Urine Indicators Related to CKD (Based on the basic specifications of urine testing combined with medical history): sg, al, su, class, htn, dm, cad, appet, pe, ane.

Regarding the data used for modelingand validation, a prior study titled "Risk Factor Prediction of Chronic Kidney Disease based on Machine Learning Algorithms" by M. A. Islam et al. (2020) was presented at the 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) in Thoothukudi, India [20]. Anotherpublic dataset for CKD validation was on Kaggle, curated by M. Jhaveri (2021) under the title "CKD Data." The dataset, accessible through the DOI: 10.24432/C5G020, contains relevant featureswhich are related to CKD diagnosis [21].

2.2. Model Selection and Construction: Decision tree algorithm was chosen considering the limited sample size (200) and self-descriptiveness.

Split Criterion. The split criterion determines the measure used to evaluate the quality of a split in the decision tree. In this case, the gain ratio was chosen as the split criterion. The gain ratio takes into account the information gain and the intrinsic information of the feature, allowing for more effective splitting.

Max Depth. The max depth parameter specifies the maximum depth or levels of the decision tree. A deeper tree can capture more complex relationships in the data, but it can also lead to overfitting. A max depth of 7 was chosen to strike a balance between capturing important patterns and avoiding overfitting.

Min Samples Split. This parameter sets the minimum number of samples required to split an internal node. A higher value helps prevent the tree from splitting too early and capturing noise or small fluctuations in the data. A value of 4 was chosen to ensure a sufficient number of samples for splitting.

Min Samples Leaf. The min samples leaf parameter specifies the minimum number of samples required to be at a leaf node. A higher value helps prevent the tree from creating small leaves based on a few samples. A value of 2 was chosen to avoid overly small leaves.

Minimal Gain. The minimal gain parameter sets the minimum improvement in the chosen split criterion required for a split to occur. It helps prevent the tree from creating splits with minimal improvement. A minimal gain of 0.1 was chosen to ensure meaningful splits.

These parameter settings were determined based on considerations of the dataset's characteristics, the desired balance between complexity and generalization, and the goal of creating an interpretable model.

3. RESULTS AND DISCUSSIONS

3.1. Performance on the Training Open Dataset

In the training phase using the open dataset, the model achieved a true positive count of 123, indicating that 123 cases of CKD were correctly identified. Additionally, it achieved a true negative count of 72, indicating that 72 non-CKD cases were correctly identified. There were no false positives (FP), meaning that no non-CKD cases were incorrectly classified as CKD. However, there were 5 false negatives (FN), indicating that 5 cases of CKD were missed by the model you can see Figure 1 and Figure 2.

3.2. Validation using the SecondOpen Dataset

In the validation phase using the second open dataset, the model achieved a true positive count of 43, indicating that 43 cases of CKD were correctly identified. Additionally, it achieved a true negative count of 115, indicating that 115 non-CKD cases were correctly identified. There were no false positives (FP), meaning that no non-CKD cases were incorrectly classified as CKD. Moreover, there were no false negatives (FN), indicating that all CKD cases were correctly identified by the model you can see Figure 2.

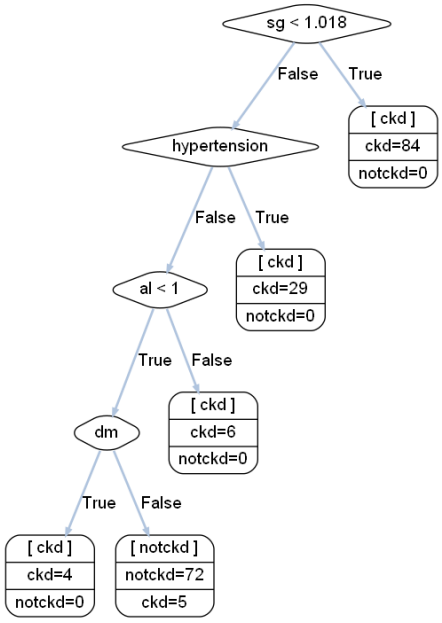

Figure 1. Performance on the Training.

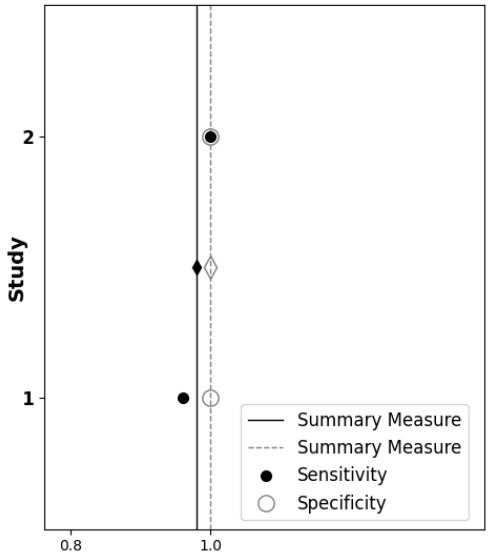

Figure 2. Trained on Study1 + Validation with Study2.

3.3. Results Analysis and Interpretation

The model's performance was evaluated on two separate datasets. The training open dataset yielded a total of 123 true positive predictions, correctly identifying patients with chronic kidney disease (CKD). Additionally, it achieved 72 true negative predictions, accurately classifying individuals without CKD. Importantly, there were no false positive predictions, indicating that the model did not falsely identify healthy individuals as having CKD. However, it did have 5 false negative predictions, meaning that 5 cases of CKD were missed by the model.

For the validation phase, the second open dataset was used to assess the model's generalization capability. The results showed promising performance, with 43 true positive predictions and 115 true negative predictions. Similar to the training dataset, there were no false positive predictions, indicating high specificity. Furthermore, there were no false negative predictions, suggesting that the model successfully identified all cases of CKD in the validation dataset.

These results demonstrate the potential of the developed model for CKD detection and classification. However, further analysis and interpretation are necessary to understand the model's limitations, generalizability, and applicability in real-world clinical settings.

3.4. The Potential and Limitations of Pure Urine Testing in CKD Diagnosis

Pure urine testing has shown significant potential as a diagnostic tool for chronic kidney disease (CKD). By analyzing specific urine indicators such as specific gravity (sg), albumin (al), and sugar (su), valuable insights can be gained regarding renal function and potential kidney damage. These non-invasive and readily available tests offer convenience and can be easily incorporated into routine clinical practices. However, it is important to note that urine testing alone may not provide a comprehensive assessment of CKD. The inclusion of additional diagnostic methods, such as blood tests and imaging studies, is crucial for a more accurate and complete evaluation of the disease.

3.5. The Effectiveness and Practicality of the Model

The developed model, based on decision tree algorithm and trained on a specific open dataset, has demonstrated promising performance in CKD detection and classification. Its ability to accurately identify true positive and true negative cases in both the training and validation datasets highlights its effectiveness. The model's self-descriptiveness, simplicity, and interpretability make it a practical choice for implementation in clinical settings with limited resources. However, it should be noted that the model's performance may vary when applied to different datasets or populations, requiring further validation and customization.

Table 2. Uppermost Recent Similar Research(July 2023)

Approach	Methodology
Early detection of CKD [22]	Systematic review discussing the use of AI for early CKD detection
CKD prediction [23]	12 machine learning-based classification algorithms, SMOTE technique, ensemble technique
Improving CKD management using AI[24]	Review article discussing the importance of ML and AI in CKD management
Automatic CKD diagnosis [25]	XGBoost classifier, Pearson correlation for feature selection, BBO algorithm for feature subset selection, SHAP analysis
Evaluating ML classification techniques [26]	Hybrid model with Gaussian Naïve Bayes, gradient boosting, decision tree as base classifiers, and random forest as meta-classifier

4. CONCLUSION

In this study, we successfully developed and validated a CKD detection model based on pure urine testing. By utilizing specific urine indicators and applying a decision tree algorithm, the model demonstrated promising performance in accurately identifying and classifying CKD cases. The inclusion of relevant clinical features and the use of a well-curated open dataset contributed to the model's effectiveness and reliability. The model's ability to leverage the simplicity and accessibility of urine testing showcases its potential as a valuable tool for early CKD diagnosis. The developed CKD detection model holds significant prospects and clinical significance. By relying on non-invasive and easily accessible urine testing, the model offers a cost-effective and convenient approach to screening and diagnosing CKD in various healthcare settings. Its potential application extends to primary care clinics, resource-limited environments, and population-based screening programs. Implementing this model can facilitate early detection and intervention, leading to improved patient outcomes, reduced healthcare costs, and better allocation of healthcare resources. While the model shows promising results, further research and validation are necessary to enhance its performance and evaluate its real-world utility. Collaboration with healthcare professionals, integration into electronic health records, and conducting prospective studies will provide valuable insights into the model's long-term predictive capabilities and clinical impact. To enhance the model's performance and broaden its applicability, several improvements and future research directions can be considered. Firstly, incorporating additional features, such as demographic information, comorbidities, and genetic markers, may improve the model's predictive accuracy. Secondly, exploring ensemble methods or other advanced machine learning algorithms could potentially enhance the model's performance by capturing more complex patterns in the data. Furthermore, conducting larger-scale studies with diverse populations and utilizing longitudinal data can provide valuable insights into the long-term predictive capabilities of the model. Lastly, integrating the model into electronic health records (EHR) systems and conducting real-world validation studies can facilitate its integration into clinical practice and assess its impact on patient outcomes. Overall, the developed model utilizing pure urine testing for CKD diagnosis shows promise but requires further investigation and refinement to optimize its effectiveness, generalizability, and real-world applicability. In conclusion, the establishment and validation of a CKD detection model based on pure urine testing offer a promising avenue for improving the early diagnosis and management of CKD. By leveraging machine learning algorithms and utilizing readily available urine indicators, this model demonstrates potential in facilitating timely interventions and improving patient outcomes.

ACKNOWLEDGEMENTS

Sincerelywould like to express my heartfelt gratitude to my supervisor, Dr. Jane Smith, for her guidance and support throughout this project. I would also like to thank my family and friends for their unwavering encouragement and understanding during this challenging time.

REFERENCES

[1] A. S. Levey and J. Coresh, "Chronic kidney disease," *Lancet*, vol. 379, no. 9811, pp. 165–180, 2012.

[2] N. R. Hill *et al.*, "Global prevalence of chronic kidney disease–a systematic review and meta-analysis," *PLoS One*, vol. 11, no. 7, p. e0158765, 2016.
[3] V. Jha *et al.*, "Chronic kidney disease: global dimension and perspectives," *Lancet*, vol. 382, no. 9888, pp. 260–272, 2013.
[4] S. Herget-Rosenthal, "Imaging techniques in the management of chronic kidney disease: current developments and future perspectives," in *Seminars in nephrology*, 2011, vol. 31, no. 3, pp. 283–290.
[5] R. T. Gansevoort *et al.*, "Chronic kidney disease and cardiovascular risk: epidemiology, mechanisms, and prevention," *Lancet*, vol. 382, no. 9889, pp. 339–352, 2013.
[6] J. R. Kuiper, K. M. O'Brien, K. K. Ferguson, and J. P. Buckley, "Urinary specific gravity measures in the US population: Implications for the adjustment of non-persistent chemical urinary biomarker data," *Environ. Int.*, vol. 156, p. 106656, 2021.
[7] S. R. Thibodeaux, *Updates in Blood Banking and Transfusion Medicine, An Issue of the Clinics in Laboratory Medicine, E-Book*, vol. 41, no. 4. Elsevier Health Sciences, 2021.
[8] G. Eknoyan *et al.*, "KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease," *Kidney int*, vol. 3, no. 1, pp. 5–14, 2013.
[9] L. A. Inker *et al.*, "KDOQI US commentary on the 2012 KDIGO clinical practice guideline for the evaluation and management of CKD," *Am. J. Kidney Dis.*, vol. 63, no. 5, pp. 713–735, 2014.
[10] K. Yamagata *et al.*, "Chronic kidney disease perspectives in Japan and the importance of urinalysis screening," *Clin. Exp. Nephrol.*, vol. 12, pp. 1–8, 2008.
[11] R. J. Glassock and A. D. Rule, "The implications of anatomical and functional changes of the aging kidney: with an emphasis on the glomeruli," *Kidney Int.*, vol. 82, no. 3, pp. 270–277, 2012.
[12] R. J. Glassock and A. D. Rule, "Aging and the kidneys: anatomy, physiology and consequences for defining chronic kidney disease," *Nephron*, vol. 134, no. 1, pp. 25–29, 2016.
[13] A. Salekin and J. Stankovic, "Detection of chronic kidney disease and selecting important predictive attributes," in *2016 IEEE International Conference on Healthcare Informatics (ICHI)*, 2016, pp. 262–270.
[14] J. C. Fink *et al.*, "Significance of serum creatinine values in new end-stage renal disease patients," *Am. J. kidney Dis.*, vol. 34, no. 4, pp. 694–701, 1999.
[15] C. Thongprayoon, W. Cheungpasitporn, and K. Kashani, "Serum creatinine level, a surrogate of muscle mass, predicts mortality in critically ill patients," *J. Thorac. Dis.*, vol. 8, no. 5, p. E305, 2016.
[16] J. Park *et al.*, "Serum creatinine level, a surrogate of muscle mass, predicts mortality in peritoneal dialysis patients," *Nephrol. Dial. Transplant.*, vol. 28, no. 8, pp. 2146–2155, 2013.
[17] S. J. Schwab, R. L. Christensen, K. Dougherty, and S. Klahr, "Quantitation of proteinuria by the use of protein-to-creatinine ratios in single urine samples," *Arch. Intern. Med.*, vol. 147, no. 5, pp. 943–944, 1987.
[18] S. D. S. Fraser and T. Blakeman, "Chronic kidney disease: identification and management in primary care," *Pragmatic Obs. Res.*, pp. 21–32, 2016.
[19] A. S. Sohal, A. S. Gangji, M. A. Crowther, and D. Treleaven, "Uremic bleeding: pathophysiology and clinical risk factors," *Thromb. Res.*, vol. 118, no. 3, pp. 417–422, 2006.
[20] M. A. Islam, S. Akter, M. S. Hossen, S. A. Keya, S. A. Tisha, and S. Hossain, "Risk factor prediction of chronic kidney disease based on machine learning algorithms," in *2020 3rd international conference on intelligent sustainable systems (ICISS)*, 2020, pp. 952–957.
[21] M. IQBAL, "Chronic KIdney Disease dataset," *https://www.kaggle.com/datasets/mansoordaku/ckdisease*, 2017. https://www.kaggle.com/mansoordaku/ckdisease.
[22] M. M. Nishat *et al.*, "A comprehensive analysis on detecting chronic kidney disease by employing machine learning algorithms," *EAI Endorsed Trans. Pervasive Heal. Technol.*, vol. 7, no. 29, pp. e1–e1, 2021.
[23] D. Chhabra, M. Juneja, and G. Chutani, "An efficient ensemble based machine learning approach for predicting Chronic Kidney Disease.," *Curr. Med. Imaging*, 2023.
[24] P. Krisanapan, S. Tangpanithandee, C. Thongprayoon, P. Pattharanitima, and W. Cheungpasitporn, "Revolutionizing Chronic Kidney Disease Management with Machine Learning and Artificial Intelligence," *Journal of Clinical Medicine*, vol. 12, no. 8. MDPI, p. 3018, 2023.
[25] M. J. Raihan, M. A.-M. Khan, S.-H. Kee, and A.-A. Nahid, "Detection of the chronic kidney disease using XGBoost classifier and explaining the influence of the attributes on the model using SHAP," *Sci. Rep.*, vol. 13, no. 1, p. 6263, 2023.
[26] H. Khalid, A. Khan, M. Zahid Khan, G. Mehmood, and M. Shuaib Qureshi, "Machine Learning Hybrid Model for the Prediction of Chronic Kidney Disease," *Comput. Intell. Neurosci.*, vol. 2023, 2023.