

Analysis and Evaluation of the Effects of Traditional Chinese Medicine Intervention on Patients with Diabetic Nephropathy in an Artificial Intelligence Environment

Ziyi Cao

The First Clinical Medical College, Shaanxi University of Chinese Medicine, Xianyang 712046, Shaanxi, China

Corresponding author: Ziyi Cao (e-mail: ziyicao299@126.com)

ABSTRACT Current evaluations of traditional Chinese medicine intervention for diabetic nephropathy often lack data-driven predictive tools, resulting in limited objectivity and inconsistent outcome assessment. To construct an objective efficacy prediction method, this study included clinical data from 2,156 patients and extracted key variables through data preprocessing and feature engineering. Logistic regression, random forest, support vector machine, and XGBoost were used for model training and comparison. The results show that the XGBoost model achieved the best predictive performance, with an AUC of 0.902. Interpretability analysis indicates that baseline urinary protein, estimated glomerular filtration rate, and traditional Chinese medicine blood stasis syndrome are important predictive factors. Subgroup analysis shows that early-stage patients without significant blood stasis syndrome obtained the best efficacy, with an effectiveness rate exceeding 80%. By integrating artificial intelligence with clinical data analysis, this study demonstrates the feasibility of quantifying traditional Chinese medicine efficacy and provides support for personalized treatment evaluation. The modeling strategy also offers references for biomedical signal analysis, intelligent diagnosis, and data-driven sensing applications.

INDEX TERMS diabetic nephropathy, traditional chinese medicine, artificial intelligence, efficacy evaluation, biomedical signal analysis

I. INTRODUCTION

DIABETIC nephropathy remains a serious microvascular complication. Its long-duration course and complex pathological mechanisms have made it a leading cause of end-stage renal disease. While the development of modern medicine has produced a core therapy based on the control of blood glucose, blood pressure and lipid ratios, and the use of renin-angiotensin system inhibitors, to delay disease progression, they work in some cases, but in others there are difficulties in their work, such as the lack of effectiveness in some cases and the continued progression of the disease. Based on the holistic principles and syndrome differentiation of traditional Chinese medicine (TCM), TCM has shown distinctive benefits and accumulated rich experience in the prevention and treatment of diabetic nephropathy by multi-target and multi-pathway intervention. However, the current assessments of TCM efficacy largely based on subjective experience of the practitioners and/or improvement in specific indicators, without a systematic tool to integrate the modern clinical indicators and TCM syndrome information to objectively and accurately predict the indi-

vidualized treatment outcome. This paper employs machine learning data mining techniques to quantitatively analyze and predict the impact of Traditional Chinese Medicine (TCM) interventions on a large-scale diabetic nephropathy population. By constructing and comparing different predictive models, a high-performance XGBoost model was developed. Furthermore, based on model interpretability methods, this paper empirically demonstrates from a data perspective that TCM syndrome elements and objective biochemical indicators are important predictive factors. Subgroup analysis further quantifies the differences in effect sizes among different patient groups, providing direct data support for the existence of a dominant population and stratification strategies in clinical practice.

The section of the article is: the current research status on the pathogenesis of diabetic nephropathy and its treatment, traditional Chinese medicine, the data sources, preprocessing, feature engineering, model construction and evaluation process; the systematic report of model performance, key predictive factor analysis, subgroup efficacy difference, and the clinical significance and limitation of this study; and

the summary of the whole article and the outlook for future research.

II. RELATED WORK

Diabetic kidney disease (DKD) is a multifaceted kidney disease triggered by a series of metabolic disorders, inflammatory responses, and fibrotic events. Its mechanistic pathology is related to the interaction of multiple dimensions (i.e., podocyte damage, immune system dysregulation, and systemic metabolic disturbances). The pathogenesis of diabetic nephropathy is multifactorial, and its essence is metabolism-inflammation-fibrosis, which was summarized by Dwivedi and Sikarwar [1]. Naaman and Bakris [2] pointed out that the combination and timely use of pillar therapies according to the patient's risk profile is the only useful clinical approach today. Barutta et al. [3] determined that hyperglycemia can cause direct and indirect damage to podocytes in a series of pathophysiological processes, leading to podocyte shedding, apoptosis, or epithelial-mesenchymal transition. Traditional Chinese medicine has accumulated rich experience in the clinical practice of preventing and treating cortical nephropathy and diabetic nephropathy. Liu et al. [4] concluded that traditional Chinese medicine has many therapeutic effects on diabetic nephropathy such as anti-inflammation, antioxidant, antifibrosis, regulating autophagy, enhancing insulin resistance, and protecting podocytes and mitochondria. Kriz et al [5] pointed out that diabetic nephropathy is a systemic disease that is gradually decreasing worldwide because it can touch all renal tissues.

Yang and Zhang[6] confirmed that innate immune dysregulation is the bridge connecting metabolic disorders and kidney damage. Azushima et al. [7] found that in diabetic patients, higher plasma lactate levels were independently associated with more severe albuminuria and decreased renal function. Sun et al . [8] systematically reviewed the key mechanisms leading to the development of diabetic nephropathy at the cellular and molecular levels . Hong et al. [9] explored the role of protein arginine methyltransferase 6 in ferroptosis and its regulatory network. Early diagnosis and accurate staging of diabetic nephropathy are crucial for improving prognosis. Luo et al. [10] pointed out that metabolomics provides a powerful biomarker discovery tool for diabetic nephropathy. These metabolite combinations (metabolic fingerprints) may be used for early risk prediction, disease classification and treatment response monitoring. This article focuses on data analysis technology in the context of artificial intelligence to conduct a systematic evaluation of the effects of traditional Chinese medicine intervention on diabetic nephropathy patients. It aims to build a efficacy prediction model through machine learning and data mining methods to make up for the current research deficiencies in quantitative assessment and personalized treatment strategies.

III. METHODS

A. DATA SOURCES AND PREPROCESSING

Electronic medical records of diabetic nephropathy patients that were either hospitalized or followed-up as outpatients in the endocrinology and nephrology departments of a tertiary hospital during the period January 2018 to December 2024 were used to gather the data of this study. Inclusion criteria were: 1) fulfilling the criterion of diagnostic type 2 diabetes;

2) fulfilling the criterion of diabetic nephropathy (persistent albuminuria greater than 30mg /24h or urine albumin/creatinine ratio greater than 30mg/g and/or estimated glomerular filtration rate somatic between 30-90mL/min/1.73m²);

3) having obtained the traditional Chinese medicine intervention (including decoctions and prepared Chinese medicine or by Among exclusion criteria are;

- 1) other serious primary kidney disease;
- 2) malignant tumours, severe heart failure (NYHA classes III-IV), active liver disease or infection;
- 3) pregnant or lactating women;
- 4) those with severe missing clinical data which is not supplementable.

A total of 2156 patients meeting the criteria were ultimately included. The collected data included:

- 1) Demographic data (age, sex, disease duration);
- 2) Baseline clinical indicators: fasting blood glucose, glycated hemoglobin, systolic blood pressure, diastolic blood pressure, serum creatinine, estimated glomerular filtration rate, 24-hour urine protein quantification or urine albumin/creatinine ratio, serum albumin, total cholesterol, triglycerides, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, etc.;
- 3) Traditional Chinese medicine syndrome information: two physicians with the title of associate chief physician or above independently diagnosed and classified the syndromes (main syndromes included Qi and Yin deficiency, Liver and Kidney Yin deficiency, Yin and Yang deficiency, combined with blood stasis, dampness, and turbid toxin syndrome, etc.), and scored the severity of main symptoms (such as fatigue, dry mouth and throat, soreness of the lower back and knees, edema, etc.) (0-3 points);
- 4) Treatment information: specific Chinese herbal formulas or proprietary Chinese medicines used, and duration of treatment;
- 5) Efficacy evaluation indicators: changes in the above clinical indicators after 3 months and 6 months of treatment. These post-treatment values were not used as input features for model training. The efficacy criteria were based on relevant guidelines, and the primary efficacy endpoint was defined as a decrease in estimated glomerular filtration rate of $\geq 15\%$ from baseline after treatment and/or a decrease in urinary protein quantification of $\geq 30\%$. Based on this criterion, patients were divided into "effective group" and "ineffective group" [11] . This study was approved by the Ethics Committee of Shaanxi University of Chinese Medicine, and was performed in accordance with the princi-

ples of the Declaration of Helsinki. All eligible participants signed an informed consent form.

Data preprocessing is a crucial step in ensuring model quality. First, missing values are handled: continuous variables with a missing value rate below 10% are imputed using the median, while variables with a missing value rate above 10% are removed. Categorical variables are imputed using the mode. Second, outlier detection is performed using box plots to identify outliers, which are then corrected or removed based on clinical significance. Finally, continuous numerical features are standardized (Z-score standardization) to eliminate the influence of units of measurement and facilitate model convergence.

B. FEATURE ENGINEERING AND DATASET CONSTRUCTION

The purpose of feature engineering is to construct a feature set that is more informative for the prediction target from the raw data. First, feature selection is performed. Based on clinical importance and univariate analysis (such as t test, chi-square test, ANOVA), baseline variables that show significant differences ($P < 0.05$) between the effective and ineffective groups are selected as candidate features. Second, considering the correlation between clinical indicators, the Pearson correlation coefficient matrix between features is calculated, and for highly correlated features (correlation coefficient > 0.8), the one with the stronger correlation to the efficacy endpoint is retained. In addition, for TCM syndrome information, in addition to directly using syndrome classification and symptom scores, composite syndrome scores, such as “Qi deficiency score”, “Yin deficiency score”, and “blood stasis score”, are also constructed to quantify the severity of syndromes [12,13]. Following the feature selection tasks, 28 features were finally incorporated in the model construction as follows age, duration of diabetes, baseline glycosylated hemoglobin, baseline systolic blood pressure, baseline estimated glomerular filtration rate, baseline urinary protein quantification, baseline serum albumin, baseline total cholesterol, presence of blood stasis syndrome, Qi deficiency score, and Yin deficiency score. The preprocessed dataset was randomly split into a training set (1509 cases) and a test set (647 cases) in the ratio of 7:3 without any bias

so as to make the ratio of the number of cases successfully processed to failure in the two groups equal to the whole dataset.

C. ARTIFICIAL INTELLIGENCE MODEL CONSTRUCTION AND EVALUATION

This study aims to construct a classification model to predict the effectiveness of traditional Chinese medicine interventions for patients. Four different types of machine learning models were trained on the training set, and their performance was compared.

Logistic Regression: As a baseline model, logistic regression has the advantages of simple model and strong interpretability. It uses the sigmoid function to map linearly combined features to probability values between 0 and 1. **Random Forest:** This is an ensemble learning algorithm that classifies data by constructing multiple decision trees and combining their voting results. It effectively handles high-dimensional features, is insensitive to outliers and noise, and has good resistance to overfitting.

Support Vector Machine (SVM): In cases where data is linearly inseparable, SVM uses a kernel function to map the data to a higher-dimensional space to achieve linear separation. This study uses radial basis functions as the kernel function.

Gradient boosting decision trees: Represented by the XGBoost algorithm, it is a sequential ensemble method in which each new tree is designed to correct the errors of the previous tree, and usually achieves high prediction accuracy.

Model training was performed on the training set using 10-fold cross-validation to optimize model hyperparameters (such as the number and maximum depth of trees in random forests, and the learning rate of XGBoost). Model performance was evaluated on an independent test set (Table 1). Evaluation metrics included accuracy, precision, recall, F1 score, and area under the receiver operating characteristic (AUC) curve. The AUC value is the core metric for measuring the model's discriminative ability; the closer the value is to 1, the better the model's performance. The optimal model was selected by comparing the performance of different models on the test set [14,15].

TABLE 1. Performance comparison of four machine learning models on the test set

Model Name	accuracy	Accuracy	Recall rate	F1 score	AUC
Logistic Regression	0.741	0.752	0.728	0.740	0.801
Random Forest	0.813	0.826	0.798	0.812	0.889
Support Vector Machine	0.768	0.781	0.753	0.767	0.842
XGBoost	0.832	0.845	0.818	0.831	0.902

Results Analysis: As shown in Table 1, among the four models, the ensemble learning methods (random forest and XGBoost) generally outperformed the traditional linear models (logistic regression) and support vector machines.

The XGBoost model performed best across all metrics, achieving an accuracy of 83.2% and an AUC of 0.902, indicating its excellent discriminative ability and ability to predict the efficacy of TCM interventions. Therefore,

subsequent analyses will be based on the XGBoost model.

IV. RESULTS AND DISCUSSION

A. KEY PREDICTOR FACTOR ANALYSIS

To gain a deeper understanding of the decision-making process of the XGBoost model and identify key factors

influencing efficacy, the SHAP method was used for model interpretability analysis. SHAP values quantify the contribution of each feature to a single prediction result; the larger the absolute value, the higher the importance of the feature. Based on the SHAP analysis of the entire test set, the feature importance ranking was obtained, as shown in Figure 1:

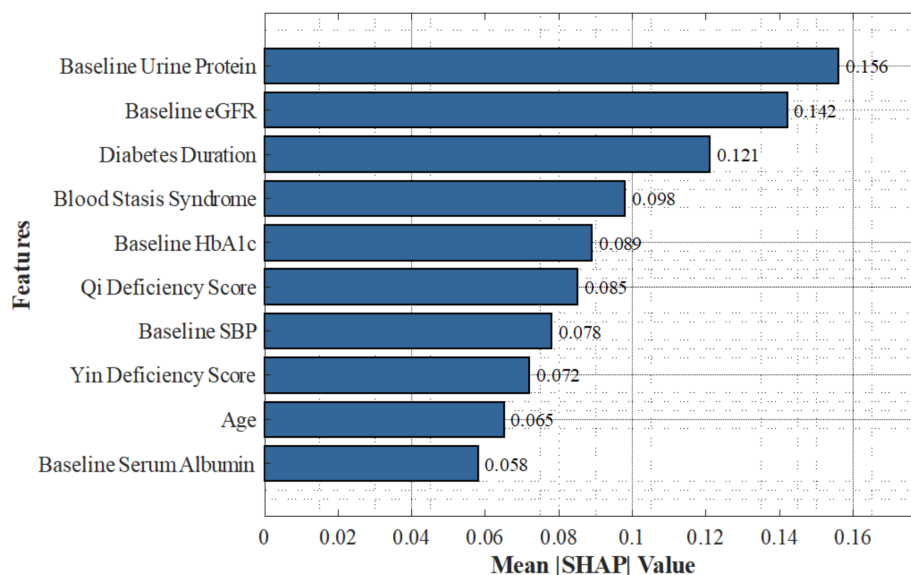


FIGURE 1. Feature Importance

The results showed that the most critical predictors of the effectiveness of TCM intervention were baseline urinary protein quantification and estimated glomerular filtration rate (GFR). This indicates that the degree of kidney damage before treatment is a crucial factor in determining efficacy. Higher urinary protein quantification and lower estimated GFR (i.e., poorer kidney function) reduce the likelihood of effective treatment, emphasizing the importance of early intervention. Long-term diabetes was also a significant negative predictor, possibly related to irreversible kidney structural damage caused by long-term hyperglycemia. Notably, the TCM syndrome factors of “whether there is concurrent blood stasis syndrome” and “qi deficiency score” were among the most important characteristics, with a negative correlation in the SHAP value; that is, the higher the concurrent blood stasis syndrome or the higher the qi deficiency score (the more severe the qi deficiency), the worse the efficacy. This data confirms the guiding value of TCM theories of “prolonged illness entering the collaterals” and “qi deficiency and blood stasis” in diabetic nephropathy, highlighting the key role of blood-activating and stasis-removing methods and qi-tonifying methods in treatment. Poor baseline glycated hemoglobin and systolic blood pressure control also negatively impacted efficacy, which is entirely consistent with the modern medical emphasis on controlling blood sugar and blood pressure.

B. PREDICTION OF EFFICACY IN DIFFERENT SUBGROUPS OF PATIENTS

Subgroup analyses of the test set patients according to significant characteristics were performed to study the predictive effectiveness of the model in the various groups of patients in order to investigate the possibility of implementing a personalized treatment plan. On the stage of stratification, it was based on the analysis of the baseline estimated glomerular filtration rate (eGFR) (stage 2 CKD stage: eGFR 60-89 ml/min/1.73m²; and eGFR 30-59 ml/min/1.73m²) and significant blood stasis syndrome (determined by blood stasis score). Each subgroup compared the predicted efficacy (i.e. the proportion of cases that the model predicts to be effective) and the observed efficacy. See Table 2:

TABLE 2. Comparison of model-predicted efficacy and actual efficacy in different patient subgroups

patient subgroup	Number of cases	Model prediction efficiency	Actual observation efficiency
Overall test set	647	58.3%	57.2%
CKD Phase 2	285	72.6%	71.9%
CKD Phase 3	362	47.2%	46.1%
Combined with significant blood stasis syndrome	254	41.7%	40.2%
No significant blood stasis syndrome	393	68.7%	67.4%
Stage 2 CKD without significant blood stasis	152	81.5%	80.9%
Stage 3 CKD with significant blood stasis syndrome	118	32.2%	30.5%

The observed efficacy rate was very comparable to the predicted efficacy rate and the model was again beneficial in terms of the uniformity that the subgroups showed. More to the point, subgroup analysis showed a high level of heterogeneity in efficacy. The rate of efficacy of treatment of the patients with early-stage diabetic nephropathy (CKD stage 2) was greater as compared to the patients with intermediate stage (CKD stage 3). There was a much higher efficacy among patients who had no national blood stasis syndrome when compared to those who had it. In the patients on CKD stage 2 and those who did not have major blood stasis syndrome the treatment efficacy rate was more than 80, which indicates that this is one group that will respond well to TCM intervention. The efficacy rate was however, much lower in patients who had already progressed to CKD stage 3 with great deal of blood stasis syndrome and was only 30.5 percent. These patients might also need more violent and efficient comprehensive treatment methods or their combination with TCM should be taken into account along with promoting the blood circulation and eliminating the blood stasis. The results have the direct data support of attaining accurate stratification and individual recommendation of treatment to diabetic nephropathy patients.

C. DISCUSSION OF THE CLINICAL EFFICACY AND LIMITATIONS OF THE MODEL

The built XGBoost prediction model showed good discriminative power and calibration with the AUC of 0.902 which is an indication of its potential clinical use value of the study in identifying patients with diabetic nephropathy that would respond to intervention with traditional Chinese medicine. The strengths of the model are: First, the model combines the objective indications of the modern medicine with TCM syndrome information, which creates a more comprehensive predictive feature system. More specifically, it establishes the significant predictive power of such TCM syndrome aspects as stasis of blood and deficit of qi, contributing to the profound unification of TCM and Western medicine assessment parameters. Second, the model has a good interpretability. SHAP analysis may visually indicate the direction and extent to which each feature affects the prediction outcomes, making clinical practitioners more familiar with the decision-making logic of the model, and not a black box, which makes the model more clinically

acceptable. Thirdly, the results of subgroup analysis can give clinicians intuitive results. Indicatively, with low probability of efficacy in a patient, the patient can be warned early, the surveillance may be intensified or treatment plan may be modified within a timely manner.

Nevertheless, there are also certain limitations of this study. First, it is a retrospective single-centre study that could be affected by the selection and information bias. The external generalization capacity of the model remains to be approved by means of multi center, prospective research. Second, the intervention regimens of TCM adopted are heterogeneous as they comprise various TCM compound prescription and prepared drugs. This can provide a confounding factor that cannot be controlled whilst the reflection is based on the real clinical practice. The next round of studies may be on particular types of syndromes or on-individual prescriptions to be studied. Lastly, the features that are present are mostly standard clinical pointers and TCM syndromes. The future study might include a more precise predictive model by adding more molecular markers using genomics, proteomics, or metabolomics, and medical imaging information.

V. CONCLUSION

This study utilizes machine learning techniques within an artificial intelligence framework to evaluate the efficacy of Traditional Chinese Medicine (TCM) intervention in the treatment of diabetic nephropathy. By constructing and testing an XGBoost predictive model, the effectiveness of personalized validation can be assessed with relatively high precision. The study found that initial renal function (urinary protein, estimated glomerular filtration rate), duration of diabetes, blood glucose and blood pressure control levels, and TCM syndrome factors such as blood stasis and qi deficiency are all important predictors of efficacy. Subgroup analysis also showed differences in efficacy among different patient groups and among those who benefited from TCM intervention. This study not only proposes a novel, data-driven approach to predict the efficacy of TCM in treating diabetic nephropathy but also interprets the factors influencing TCM efficacy through interpretability analysis, thus providing valuable information for developing individualized TCM treatment plans in clinical practice. In the future, through multi-center prospective data validation and

multi-dimensional integration, these artificial intelligence models will become powerful tools for clinicians to achieve precision medicine, much like the meticulous calibration performed by workers to ensure the perfect operation of complex machines.

AUTHOR CONTRIBUTIONS

Ziyi Cao designed, collected and analyzed the data, and drafted the manuscript. Ziyi Cao conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Ziyi Cao participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

HUMAN RESEARCH SUBJECTS

This study was approved by the Ethics Committee of Shaanxi University of Chinese Medicine, and was performed in accordance with the principles of the Declaration of Helsinki. All eligible participants signed an informed consent form.

REFERENCES

- [1] S. Dwivedi and M. S. Sikarwar, "Diabetic nephropathy: pathogenesis, mechanisms, and therapeutic strategies," *Hormone and Metabolic Research*, vol. 57, no. 01, pp. 7-17, 2025, doi: 10.1055/a-2435-8264.
- [2] S. C. Naaman and G. L. Bakris, "Diabetic nephropathy: update on pillars of therapy slowing progression," *Diabetes Care*, vol. 46, no. 9, pp. 1574-1586, 2023, doi: 10.2337/dci23-0030.
- [3] F. Barutta, S. Bellini, and G. Gruden, "Mechanisms of podocyte injury and implications for diabetic nephropathy," *Clinical Science*, vol. 136, no. 7, pp. 493-520, 2022, doi: 10.1042/CS20210625.
- [4] X. J. Liu, X. K. Hu, H. Yang, et al., "A review of traditional Chinese medicine on treatment of diabetic nephropathy and the involved mechanisms," *The American journal of Chinese medicine*, vol. 50, no. 07, pp. 1739-1779, 2022, doi: 10.1142/S0192415X22500744.
- [5] W. Kriz, J. Löwen, and H. J. Gröne, "The complex pathology of diabetic nephropathy in humans," *Nephrology Dialysis Transplantation*, vol. 38, no. 10, pp. 2109-2119, 2023, doi: 10.1093/ndt/gfad052.
- [6] M. Yang and C. Zhang, "The role of innate immunity in diabetic nephropathy and their therapeutic consequences," *Journal of Pharmaceutical Analysis*, vol. 14, no. 1, pp. 39-51, 2024, doi: 10.1016/j.jpha.2023.09.003.
- [7] K. Azushima, J. P. Kovalik, T. Yamaji, et al., "Abnormal lactate metabolism is linked to albuminuria and kidney injury in diabetic nephropathy," *Kidney International*, vol. 104, no. 6, pp. 1135-1149, 2023, doi: 10.1016/j.kint.2023.08.006.
- [8] T. Sun, Y. Guo, Y. Su, et al., "Molecular mechanisms of diabetic nephropathy: A narrative review," *Cell Biology International*, vol. 48, no. 9, pp. 1240-1253, 2024, doi: 10.1002/cbin.12212.
- [9] K. Omote, T. Aikawa, D. Sunaga, et al., "Bright kidney: a sign of contrast induced nephropathy after percutaneous coronary intervention," *Oxford Medical Case Reports*, vol. 2025, no. 7, 2025, doi: 10.1093/omcr/omaf110.
- [10] Y. Luo, W. Zhang, and G. Qin, "Metabolomics in diabetic nephropathy: Unveiling novel biomarkers for diagnosis," *Molecular Medicine Reports*, vol. 30, no. 3, pp. 1-13, 2024, doi: 10.3892/mmr.2024.13280.
- [11] B. L. Abhirami, A. A. Krishna, A. Kumaran, et al., "Targeting NF- κ B in diabetic nephropathy: exploring the therapeutic potential of phytoconstituents," *Archives of Pharmacol Research*, vol. 48, no. 7, pp. 577-637, 2025, doi: 10.1007/s12272-025-01555-z.
- [12] W. Zhang, M. Liu, L. Liu, et al., "Polarization-Enhanced Janus MoSSe Plasmonic Biosensor for Accurate Diabetic Nephropathy Diagnosis," *Nano Letters*, vol. 25, no. 34, pp. 13031-13039, 2025, doi: 10.1021/acs.nanolett.5c03392.
- [13] T. Hu, J. Yue, Q. Tang, et al., "The effect of quercetin on diabetic nephropathy (DN): a systematic review and metaanalysis of animal studies," *Food & function*, vol. 13, no. 9, pp. 4789-4803, 2022, doi: 10.1039/D1FO03958J.
- [14] M. Baraka, A. Khalil, and N. Okasha, "Efficacy of Colchicine Preloading in Mitigating Contrast-Induced Nephropathy in Diabetic Patients Undergoing Percutaneous Coronary Intervention," *Egyptian Journal of Hospital Medicine*, vol. 98, no. 1, 2025, doi: 10.21608/ejhm.2025.416971.
- [15] S. Mutter, E. Valo, V. Aittomäki, et al., "Urinary metabolite profiling and risk of progression of diabetic nephropathy in 2670 individuals with type 1 diabetes," *Diabetologia*, vol. 65, no. 1, pp. 140-149, 2022, doi: 10.1007/s00125-021-05584-3.