Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Development and Internal Validation of an Interpretable Machine-Learning Model for Predicting 6-Month eGFR- Category Decline in Adults With Diabetes and Renal Dysfunction

Yanru Liu
Shanghai pudong Hospital
Tao Yang
Shanghai pudong Hospital
Yi Lu
Shanghai pudong Hospital
Weifu Liang
Fudan University
Jingru Gong
Shanghai pudong Hospital
Huiping Lu
Luhuiping310@163.com
Shanghai pudong Hospital

Article

Keywords: diabetes mellitus, renal dysfunction, kidney function decline, machine learning, CatBoost, prediction model, Chinese patent medicine

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11503634/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Development and Internal Validation of an Interpretable Machine-Learning Model for Predicting 6-Month eGFR-Category Decline in Adults With Diabetes and Renal Dysfunction

Yanru Liu [1,#], Tao Yang [1,2,#], Yi Lu [1], Weifu Liang [3], Jingru Gong [1,*], Huiping Lu [1,*]

[1] Department of Pharmacy, Shanghai Pudong Hospital, Fudan University Pudong Medical Center, 201399 Shanghai, China

[2] Intelligent Medicine Institute, Shanghai Medical College, Fudan University, 200032 Shanghai, China

[3] Information Technology Department, Shanghai Pudong Hospital, Fudan University Pudong Medical Center, 201399 Shanghai, China

* Correspondence: luhuiping310@163.com (H.L.); 261911922@qq.com (J.G.)

[#] These authors contributed equally to this work.

Abstract

Background: Adults with diabetes and renal dysfunction are at increased risk of short-term kidney function decline, yet individualized risk prediction remains challenging in routine practice. We developed and internally validated an interpretable machine-learning model for 6-month eGFR-category decline and conducted exploratory analyses of associations between Chinese patent medicine (CPM) exposure and this outcome.

Methods: This single-center retrospective cohort study included adults with diabetes and renal dysfunction who visited Shanghai Pudong Hospital between January 2021 and December 2024. The outcome was progression to a more advanced eGFR category approximately 6 months after the index date. Candidate predictors comprised demographic characteristics, laboratory indicators, comorbidity burden, conventional medications, SGLT2 inhibitor use, and four CPMs. Nine algorithms were compared using 10-fold cross-validation in the development set, and the selected model was evaluated in a randomly assigned internal test set.

Results: Of 2,022 patients (1,617 in the development set and 405 in the internal test set), 240 (11.87%) experienced 6-month eGFR-category decline. CatBoost achieved the highest mean cross-validated AUROC (0.89, 95% CI 0.86–0.91) and an AUROC of 0.91 in the internal test set. In a post hoc resource-prioritization analysis, the top 5% of predicted risks (n = 21) included 19 events, corresponding to a sensitivity of 39.58%, specificity of 99.44%, and positive predictive value of 90.48%. Exploratory DML analyses suggested a negative adjusted association for SSN exposure, whereas estimates for the other CPMs were imprecise and crossed the null.

Conclusions: An interpretable CatBoost model using 15 predictors showed good internal discrimination for 6-month eGFR-category decline. However, the post hoc top-5% rule identified fewer than half of all events. Calibration assessment, complete uncertainty reporting, temporal or external validation, and prospective impact evaluation are required before clinical use. CPM findings are exploratory, hypothesis-generating associations and should not be interpreted as evidence for treatment selection.

Keywords: diabetes mellitus; renal dysfunction; kidney function decline; machine learning; CatBoost; prediction model; Chinese patent medicine

1. Introduction

Diabetes is a leading cause of kidney function impairment worldwide, and the coexistence of diabetes and renal dysfunction substantially increases the risks of kidney failure, cardiovascular events, mortality, and healthcare expenditure [1]. Although diabetic kidney disease (DKD) is the usual clinical framing, its rigorous diagnosis requires comprehensive clinical adjudication, longitudinal assessment of albuminuria and estimated glomerular filtration rate (eGFR), exclusion of alternative kidney diseases, and, in selected cases, histological confirmation [2]. In routine hospital information systems, these diagnostic elements are frequently incomplete [3]. For real-world prediction studies, therefore, focusing on adults with diabetes and renal dysfunction offers a more pragmatic and reproducible framework than assuming a gold-standard diagnosis of DKD.

Current guidelines emphasize early identification of high-risk patients and comprehensive cardiorenal risk reduction guided by eGFR, albuminuria, metabolic control, blood pressure, and comorbidities [2,4,5]. Sodium–glucose cotransporter 2 inhibitors (SGLT2i) have become foundational therapy for patients with diabetes and chronic kidney disease (CKD), supported by landmark trials demonstrating reduced risks of kidney disease progression and cardiovascular events [6–8]. Glucagon-like peptide-1 receptor agonists and nonsteroidal mineralocorticoid receptor antagonists have further expanded the therapeutic armamentarium [4,9]. Nevertheless, many patients continue to experience progressive kidney function decline despite guideline-directed care [7], underscoring the need for individualized risk stratification to inform follow-up and management planning.

In China, Chinese patent medicines (CPMs) are widely used as adjunctive therapies for kidney disease and diabetes-related renal complications. Preparations such as Shenshuaining (SSN), Keluoxin (KLX), Jinshuibao (JSB), and Bailing capsules (BLJN) are commonly prescribed, and randomized trials,

systematic reviews, and evidence-mapping studies have reported possible benefits for proteinuria, renal function, symptoms, or quality of life [10–12]. However, this evidence is heterogeneous, and observational comparisons within contemporary SGLT2i-based care remain vulnerable to confounding by indication, exposure misclassification, and limited treatment overlap. These limitations preclude treatment-selection claims derived from a prognostic model.

Machine learning can integrate routine clinical, laboratory, and medication data to capture nonlinear relationships and identify patients at high short-term risk [13–15]. Feature-importance and Shapley additive explanations (SHAP) analyses can improve interpretability by linking predictions to clinically familiar factors [14,15]. Double machine learning can be used as an exploratory adjustment framework for treatment–outcome associations in observational data, but its estimates remain dependent on causal assumptions and cannot replace randomized evidence [16].

Therefore, this study aimed to develop and internally validate an interpretable machine-learning model for 6-month eGFR-category decline in adults with diabetes and renal dysfunction. We compared nine algorithms, interpreted the selected model with SHAP, and implemented a research prototype for individualized risk estimation. Post hoc clustering, subgroup comparisons, and CPM exposure–outcome analyses were treated as exploratory and non-causal.

2. Materials and Methods

2.1. Study Design and Population

This single-center retrospective cohort study was designed to develop, internally validate, interpret, and implement a machine-learning model for predicting 6-month eGFR-category decline in adults with diabetes and renal dysfunction using routinely collected clinical data from Shanghai Pudong Hospital. A baseline prediction design was adopted in which each patient contributed one index record. The study is reported in accordance with the TRIPOD+AI statement [17].

Electronic medical records were retrieved from the hospital information system, laboratory information system (LIS), and nursing information system for the period from January 1, 2021 to December 31, 2024. Because routine hospital data do not always contain all elements required for gold-standard adjudication of DKD, the target population was defined as adults with diabetes and renal dysfunction rather than patients with comprehensively or pathologically adjudicated DKD.

Eligible patients were adults aged ≥18 years with a recorded diagnosis of diabetes and evidence of kidney abnormality on or before the index date. Kidney abnormality was operationally defined as at least one of the following: eGFR <60 mL/min/1.73 m², urinary albumin-to-creatinine ratio (UACR) ≥30 mg/g, documented quantitative or qualitative proteinuria above the laboratory reference limit, or an electronic diagnosis of CKD or DKD. Because a single abnormal measurement or diagnosis code does not establish chronicity or etiology, the population is described as adults with diabetes and renal dysfunction. Patients were additionally required to have at least three eGFR measurements during the study period. Exclusion criteria were severely incomplete clinical data, polycystic kidney disease or another primary kidney disease requiring active immunosuppressive therapy, use of immunosuppressive agents or glucocorticoids within the preceding 3 months, pregnancy, and lactation.

2.2. Index Date and Candidate Predictors

One baseline prediction record was constructed for each eligible patient. The index date was defined as the date of the earliest qualifying eGFR measurement on which the patient simultaneously met all eligibility criteria; this index eGFR defined the baseline kidney-function category. A candidate index date was eligible only when sufficient pre-index records were available to ascertain the prespecified 180-day laboratory and 90-day medication-exposure windows and when an eligible follow-up eGFR measurement was available 150–210 days after the index date for outcome assessment. When more than one date met these conditions, the earliest was selected, so that each patient contributed only one index record. All predictors were ascertained on or before the index date; post-index laboratory results, prescriptions, and treatment changes were not used as model inputs.

Candidate predictors comprised demographic characteristics, laboratory indicators, comorbidity burden (Charlson Comorbidity Index), and medication exposure. Demographic variables were age at the index date and sex.

Laboratory variables were obtained from the LIS. For each indicator, the measurement closest to the index date within the 180 days before or on the index date was used as the baseline value; if no measurement was available within this window, the value was treated as missing and handled during preprocessing. Laboratory variables covered renal function, urinary protein burden, glycemic control, lipid metabolism, inflammation, liver function, nutritional status, and hematologic status, and included

eGFR, serum creatinine, UACR, urine protein, glycated hemoglobin (HbA1c), triglycerides, high-sensitivity C-reactive protein (hsCRP), red blood cell (RBC) count, white blood cell (WBC) count, hemoglobin, albumin, and total bilirubin, among others.

Medication exposure was defined from prescription records within the 90 days before or on the index date; patients were considered exposed if they had at least one prescription during this window or an active prescription overlapping the index date. Medication variables comprised conventional antidiabetic, antihypertensive, lipid-lowering, antiplatelet, and kidney-protective agents, as well as the four CPMs of interest (SSN, KLX, JSB, and BLJN), each encoded as a binary indicator. Definitions, abbreviations, and coding schemes for all candidate predictors are provided in Supplementary Table S2.

2.3. Outcome Definition and Data Collection

The primary outcome was progression to a more advanced (lower) eGFR category at approximately 6 months after the index date. Baseline and follow-up categories were defined using the KDIGO G1–G5 thresholds: G1, ≥90; G2, 60–89; G3a, 45–59; G3b, 30–44; G4, 15–29; and G5, <15 mL/min/1.73 m² [2]. The follow-up value was the eligible eGFR measurement closest to day 180 within a prespecified window of 150–210 days after the index date; when two measurements were equally close to day 180, the later measurement was selected. Measurements obtained during a documented acute kidney injury (AKI) episode, identified from an AKI diagnosis or a KDIGO-consistent acute creatinine change when sufficient creatinine data were available, were excluded. Patients without an eligible follow-up eGFR within this window were not included in the analytic cohort. The outcome was coded as 1 when follow-up eGFR fell into a more advanced G category than index eGFR (for example, G2 to G3a) and as 0 otherwise. Because category worsening was not required to persist for ≥90 days, the outcome is termed 6-month eGFR-category decline rather than confirmed CKD progression. eGFR was used as reported by the LIS, which calculated it from serum creatinine using the 2021 CKD-EPI creatinine equation [18], and was not recalculated for this analysis.

All candidate predictors and outcomes were extracted from structured electronic records, including demographic information, diagnoses, laboratory results, and medication orders. Data were checked for logical consistency, duplicate records, biologically implausible values, and missingness before model development. Patient identifiers were removed before analysis.

2.4. Data Preprocessing and Feature Selection

Candidate variables were screened for missingness and distributional imbalance before modeling. Fasting insulin (31.26%) and fasting C-peptide (19.68%) had the highest missingness (Table S1); all other candidate variables had lower missingness. Two extremely imbalanced medication variables, finerenone and dorzagliatin, were excluded, leaving 40 non-CPM variables and four prespecified CPM indicators for feature engineering. The CPM indicators were retained for the prespecified exploratory medication-scenario component; their inclusion should not be interpreted as evidence that they improved prognostic performance.

The analytic cohort was stratified by outcome status and randomly divided at the patient level into a development set (n = 1,617) and an internal test set (n = 405) in an 8:2 ratio. All preprocessing steps—random forest–based imputation [19], standardization of continuous variables where required, one-hot encoding of binary variables, and ordinal encoding of multicategory variables—were fitted on development-set data only and then applied to the internal test set; the test-set imputer used development-set parameters exclusively. Within cross-validation, preprocessing and the feature-selection procedure were refitted inside each training fold, so that no test-set or held-out-fold information contributed to imputation, model training, hyperparameter tuning, or feature selection.

Feature selection used an XGBoost-based stepwise procedure. XGBoost importance scores were first estimated for all candidate predictors, which were then ranked by importance, and stepwise forward selection was used to identify a compact subset with the highest cross-validated discrimination. Eleven non-CPM variables were selected: eGFR, SGLT2i use, RBC count, hsCRP, UACR, age, WBC count, HbA1c, triglycerides, sex, and total bilirubin. The four prespecified CPM indicators (SSN, KLX, JSB, and BLJN) were retained for exploratory medication-scenario analyses, yielding 15 predictors in the implemented model.

2.5. Model Development and Interpretation

Nine candidate algorithms were compared: CatBoost, XGBoost, decision tree, support vector machine (SVM), FTTransformer, TabulaRNN, recurrent neural network (RNN), long short-term memory network (LSTM), and temporal convolutional network (TCN). Because the dataset contained one baseline record per patient rather than longitudinal sequences, the sequence-oriented neural networks were included only as exploratory benchmarks; tree-based and other tabular-data models were considered the primary candidates.

Hyperparameters were tuned in the development set using grid search with 10-fold cross-validation. The algorithm with the highest mean cross-validated AUROC and favorable stability across folds was selected as the final model. CatBoost [20] was selected on this basis (Section 3.4); its final settings were iterations = 500, early_stopping_rounds = 100, learning_rate = 0.03, and depth = 5. The internal test set was reserved for subsequent performance evaluation and was not used for model selection.

Model interpretability was assessed using feature-importance analysis and SHAP values [21]. Global feature importance ranked the contribution of each predictor, and SHAP summary and dependence plots visualized the direction and magnitude of feature effects.

2.6. Exploratory Clustering, Subgroup, and Medication-Association Analyses

Post hoc subgroup comparisons of CatBoost-predicted probabilities were conducted according to SGLT2i use, sex, and use of SSN, KLX, JSB, and BLJN. Because the medication indicators were themselves included among model inputs and the comparisons involved predicted probabilities rather than observed treatment effects, these analyses were interpreted only as descriptive comparisons of model-defined baseline risk. P values were exploratory and were not adjusted for multiple comparisons.

Post hoc K-means clustering of the selected predictors was performed in the internal test set to explore phenotypic heterogeneity, with three clusters specified and visualized by principal component analysis. Because clustering was performed after model evaluation in the same test set, it was considered an exploratory phenotyping analysis and did not provide additional validation of predictive performance.

Double machine learning (DML) was used as an exploratory adjustment framework to estimate associations between each baseline CPM exposure and the binary 6-month outcome. The target estimand was expressed on the absolute risk-difference scale. DML uses flexible nuisance modeling, orthogonalization, and sample splitting to reduce regularization bias [22]; however, because this observational analysis could not establish exchangeability, adequate treatment overlap, or absence of unmeasured confounding, the resulting estimates were interpreted as exploratory adjusted exposure–outcome associations rather than causal treatment effects or evidence for treatment recommendations [23].

2.7. Research-Prototype Implementation

The final CatBoost model was implemented as a mini-program–based research prototype. Users enter the routinely available predictors included in the final model and obtain an estimated probability of short-term eGFR-category decline. The prototype does not collect direct patient identifiers.

The prototype also allows medication indicators to be modified for non-causal scenario visualization. Because medication coefficients in a prediction model do not estimate individual treatment effects, differences between scenarios must not be interpreted as the expected benefit or harm of starting, stopping, or switching a medication. The output is intended for research demonstration and risk communication, not for autonomous clinical decision-making or treatment selection.

The interface uses structured predictor entry and returns estimated risks in an interpretable format. Clinical usability, calibration in new patients, workflow integration, safety, and impact on patient outcomes were not evaluated in this retrospective study. Accordingly, the prototype should be regarded as a research demonstration rather than a clinical decision-support tool.

2.8. Statistical Analysis

Continuous variables were summarized as median (interquartile range) because their distributions were generally skewed, and categorical variables were summarized as frequency (percentage). Between-group comparisons used the Mann–Whitney U test for continuous variables and the chi-square test or Fisher's exact test for categorical variables, as appropriate. Balance between the development and internal test sets was assessed using standardized mean differences. All P values were two-sided.

Discrimination was assessed primarily using the area under the receiver operating characteristic curve (AUROC). Cross-validated AUROC was summarized across the 10 development-set folds. In the internal test set, a post hoc resource-prioritization analysis defined a high-risk stratum as the top 5% of predicted probabilities (n = 21); sensitivity, specificity, PPV, NPV, accuracy, and F1 score were calculated from the resulting 2 × 2 classification table. Because this cutoff was rank-based within the test set rather than a prespecified clinical probability threshold, these classification measures are descriptive and should not be interpreted as a transportable clinical cutoff. Decision curve analysis was used to estimate net benefit across threshold probabilities [24].

All analyses were performed in Python 3.10.9 using scikit-learn 1.7.2, XGBoost 3.2.0, CatBoost 1.2.10, missForest 4.2.3, SHAP 0.48.0, and PyTorch 2.9.1. A fixed random seed of 1314 was used for the stratified development–test split and for cross-validation to ensure reproducibility.

3. Results

3.1. Patient Selection and Model-Development Workflow

A total of 22,568 patients who visited Shanghai Pudong Hospital between January 1, 2021 and December 31, 2024 were screened. After excluding patients younger than 18 years, 22,487 adults remained; of these, 2,317 had diabetes and renal dysfunction, and 2,058 had at least three eGFR measurements and were assessed for eligibility. Thirty-six patients were excluded (10 with polycystic kidney disease or another primary kidney disease requiring active immunosuppression and 26 with recent use of immunosuppressive agents or glucocorticoids); no patient was excluded for severely incomplete clinical data, pregnancy, or lactation.

The final analytic cohort comprised 2,022 patients, of whom 240 (11.87%) experienced 6-month eGFR-category decline; 40 of the 240 events (16.7%) were G1-to-G2 transitions. The cohort was stratified by outcome status and randomly divided in an 8:2 ratio into a development set (n = 1,617; 192 events, 11.87%) and an internal test set (n = 405; 48 events, 11.85%). The development set was used for preprocessing, imputation, cross-validation, and model development, whereas the internal test set was reserved for performance evaluation. The final model was subsequently implemented as a mini-program for individualized risk estimation and non-causal medication-scenario visualization (Figure 1).

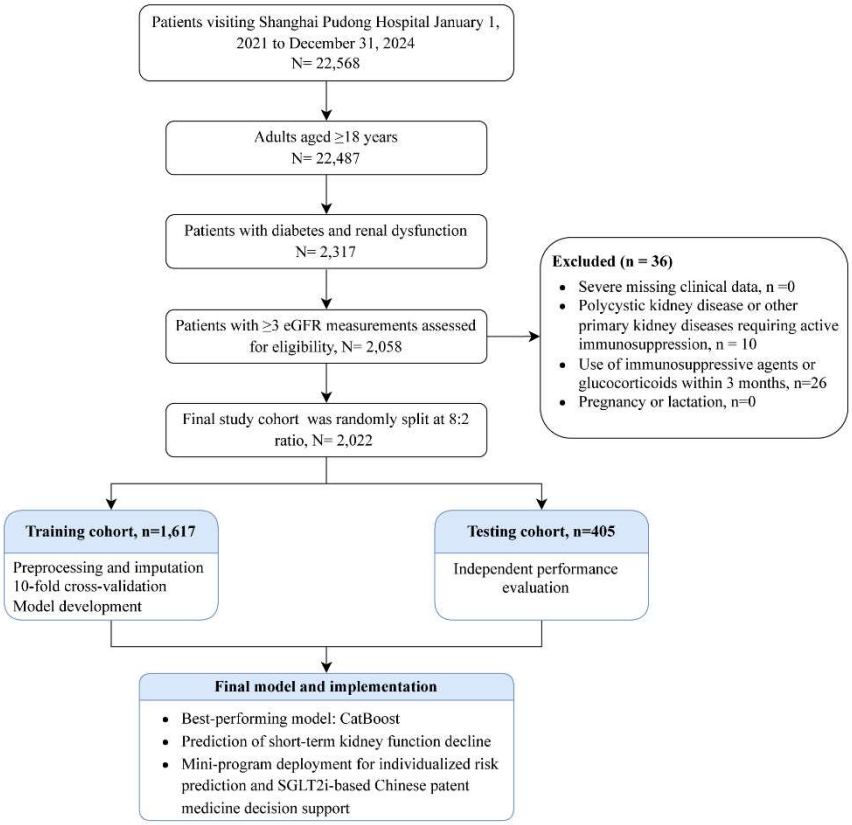


Figure 1. Study flowchart and model-development workflow. A total of 22,568 patients visiting Shanghai Pudong Hospital from January 1, 2021 to December 31, 2024 were screened; after application of the inclusion and exclusion criteria, 2,022 patients formed the analytic cohort. The cohort was stratified by outcome status and randomly split 8:2 into a development set and an internal test set. The development set was used for preprocessing, imputation, cross-validation, and model development, whereas the internal test set was reserved for performance evaluation. The final CatBoost model was implemented in a mini-program for individualized risk estimation and non-causal medication-scenario visualization. eGFR, estimated glomerular filtration rate; SGLT2i, sodium–glucose cotransporter 2 inhibitor.

Baseline characteristics of the 15 final predictors are summarized in Table 1. Patients who experienced the outcome were older and had lower eGFR, higher UACR, lower total bilirubin, lower RBC count, and less frequent SGLT2i use than those who did not. Sex, HbA1c, triglycerides, hsCRP, WBC count, and CPM exposure did not differ significantly between groups. The remaining candidate variables, predictor definitions, and development–test set balance are shown in Supplementary **Tables S1–S3**.

Table 1. Baseline characteristics of the final model predictors stratified by 6-month eGFR-category decline.

Variable	Overall (n=2022)	Non-progression (n=1782)	Progression (n=240)	P value	Missing rate %
Age (years)	60.00 (49.00-71.00)	59.00 (47.00-70.00)	69.00 (62.00-76.00)	<0.001	0.00
Male sex, n (%)	1038 (51.34)	907 (50.90)	131 (54.58)	0.316	0.00
eGFR (mL/min/1.73 m²)	99.89 (72.56-110.34)	101.73 (75.46-111.70)	90.78 (60.90-95.80)	<0.001	0.00
HbA1c (%)	7.40 (6.50-9.20)	7.30 (6.50-9.20)	7.64 (6.57-9.00)	0.277	3.51
UACR (mg/g)	30.00 (30.00-80.00)	30.00 (30.00-80.00)	30.00 (30.00-150.00)	<0.001	4.45
TBIL (μmol/L)	11.90 (9.20-15.67)	12.00 (9.30-15.70)	11.15 (8.00-15.10)	0.010	0.35
TG (mmol/L)	1.46 (1.03-2.17)	1.47 (1.03-2.18)	1.42 (1.01-2.09)	0.436	2.92
hsCRP (mg/L)	0.94 (0.50-3.14)	0.99 (0.50-3.08)	0.66 (0.50-3.26)	0.208	3.21
WBC (×10⁹/L)	6.87 (5.77-8.28)	6.88 (5.77-8.28)	6.79 (5.78-8.23)	0.432	1.73
RBC (×10¹²/L)	4.64 (4.23-5.04)	4.67 (4.29-5.09)	4.30 (3.74-4.75)	<0.001	1.73
SGLT2i, n (%)	313 (15.48)	294 (16.50)	19 (7.92)	<0.001	0.00
SSN, n (%)	63 (3.12)	54 (3.03)	9 (3.75)	0.686	0.00
KLX, n (%)	153 (7.57)	139 (7.80)	14 (5.83)	0.341	0.00
JSB, n (%)	49 (2.42)	41 (2.30)	8 (3.33)	0.451	0.00
BLJN, n (%)	104 (5.14)	89 (4.99)	15 (6.25)	0.502	0.00

Continuous variables are presented as median (interquartile range) and categorical variables as n (%). Between-group comparisons used the Mann–Whitney U test for continuous variables and the chi-square or Fisher's exact test

for categorical variables, as appropriate. Missing rates were calculated before imputation. The outcome was progression to a more advanced eGFR category at approximately 6 months. eGFR, estimated glomerular filtration rate; UACR, urinary albumin-to-creatinine ratio; TBIL, total bilirubin; TG, triglycerides; hsCRP, high-sensitivity C-reactive protein; WBC, white blood cell count; RBC, red blood cell count; SGLT2i, sodium–glucose cotransporter 2 inhibitor; SSN, Shenshuaining capsule; KLX, Keluoxin capsule; JSB, Jinshuibao capsule; BLJN, Bailing capsule.

3.2. Feature Selection Based on Importance and Incremental AUROC

Candidate predictors were ranked by XGBoost feature importance and added sequentially to evaluate changes in cross-validated discrimination. As shown in Figure 2, eGFR had the highest importance score, followed by SGLT2i use, RBC count, hsCRP, UACR, age, WBC count, HbA1c, triglycerides, sex, and total bilirubin. The mean AUROC increased rapidly with the inclusion of the top-ranked variables and then plateaued, indicating limited additional predictive gain from lower-ranked features.

On the basis of feature importance and incremental AUROC, these 11 non-CPM variables were selected as a compact predictive subset. The four prespecified CPM indicators were retained for the exploratory medication-scenario component, yielding 15 predictors in the implemented model.

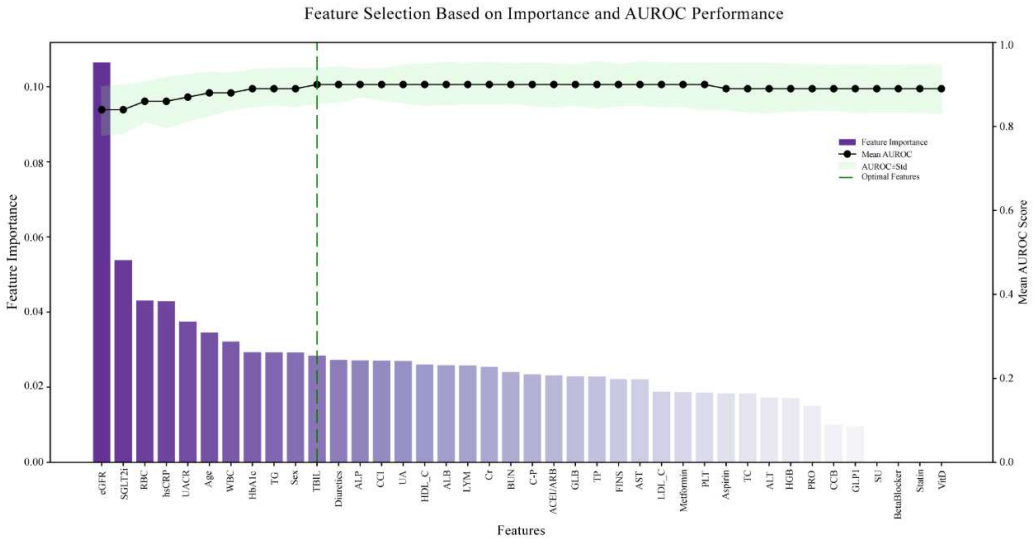


Figure 2. Feature selection based on importance and incremental AUROC. Candidate predictors were ranked according to feature importance scores. The bar plot shows the relative importance of each feature, and the line plot shows the mean AUROC obtained by sequentially adding features according to their importance ranking. The shaded area represents the standard deviation of the AUROC. The vertical dashed line indicates the optimal feature subset selected for model development. AUROC, area under the receiver operating characteristic curve; eGFR, estimated glomerular filtration rate; SGLT2i, sodium–glucose cotransporter 2 inhibitor; hsCRP, high-sensitivity C-reactive protein; UACR, urinary albumin-to-creatinine ratio.

3.3. Model Interpretation and Exploratory Cluster-Based Risk Profiling

The final CatBoost model was interpreted using SHAP analysis; unsupervised clustering was performed as a separate post hoc exploratory analysis. As shown in Figure 3A, eGFR, RBC count, HbA1c, UACR, age, hsCRP, WBC count, total bilirubin, SGLT2i use, and triglycerides were the leading contributors to model output, representing renal function, hematologic status, glycemic control, urinary protein burden, inflammation, and metabolic status.

To explore patient heterogeneity, K-means clustering was performed on the selected core predictors in the internal test set, which was partitioned into three clusters with distinct profiles (Figure 3B). Predicted probability, eGFR, RBC count, and HbA1c differed across clusters (Figure 3C–F). Cluster 1 had the lowest predicted risk and relatively higher eGFR and RBC count, whereas clusters 0 and 2 had higher predicted risk and less favorable renal or hematologic profiles. These post hoc findings are descriptive and hypothesis-generating; they do not constitute additional validation of predictive performance.

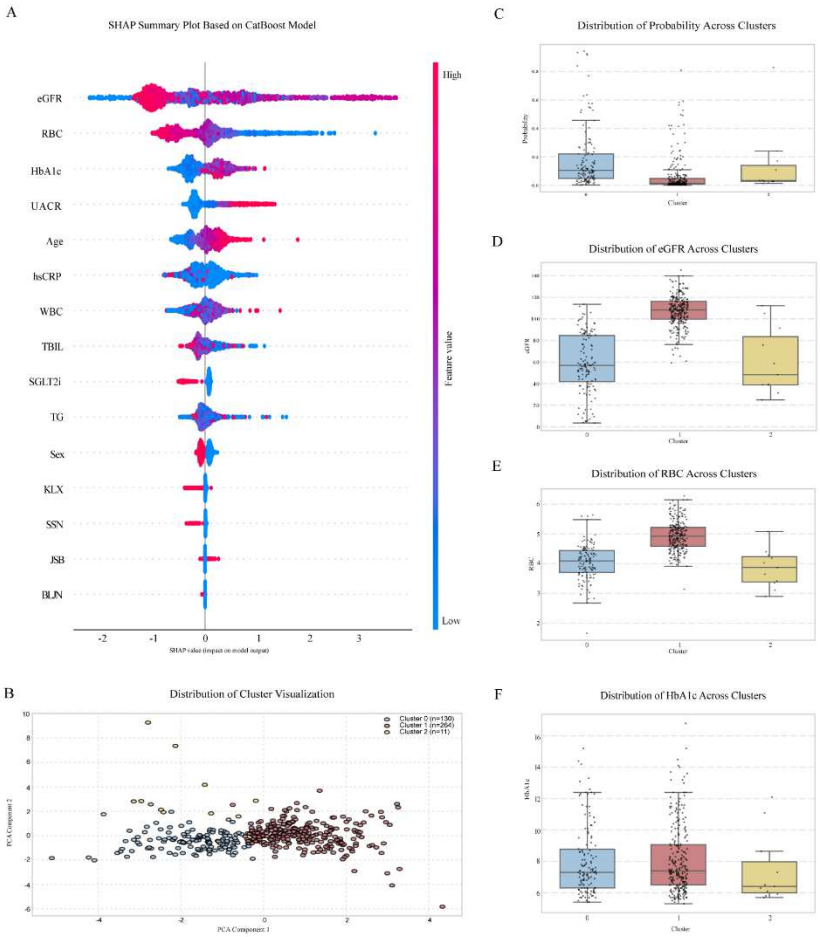


Figure 3. Model interpretation and cluster-based risk profiling using the final CatBoost model. (A) SHAP summary plot showing the contribution of each predictor to the CatBoost model output. Each dot represents an individual patient, and the color indicates the feature value from low to high. (B) K-means clustering visualization based on the

selected core predictors using principal component analysis. (C–F) Distribution of predicted probability, eGFR, red blood cell count, and HbA1c across the three clusters. Boxplots show the median and interquartile range, with dots representing individual patients. SHAP, Shapley additive explanations; eGFR, estimated glomerular filtration rate; RBC, red blood cell count; HbA1c, glycated hemoglobin; UACR, urinary albumin-to-creatinine ratio; hsCRP, high-sensitivity C-reactive protein; WBC, white blood cell count; TBIL, total bilirubin; SGLT2i, sodium–glucose co-transporter 2 inhibitor; TG, triglycerides.

3.4. Model Performance in the Internal Test Set

In the development set, CatBoost had the highest mean 10-fold cross-validated AUROC (0.89, 95% CI 0.86–0.91) and was therefore selected for evaluation in the internal test set. In the internal test set, CatBoost achieved an AUROC of 0.91 (Table 2). In a post hoc top-5% resource-prioritization analysis, 21 patients were classified as high risk, of whom 19 experienced the outcome (19 true positives, 2 false positives, 29 false negatives, and 355 true negatives). This corresponded to a sensitivity of 39.58%, specificity of 99.44%, PPV of 90.48%, NPV of 92.45%, accuracy of 92.35%, and F1 score of 0.551. Thus, this rank-based rule identified a small subgroup with a high event proportion but missed more than half of all outcome events and should not be interpreted as a clinically validated probability threshold.

CatBoost had the highest test-set AUROC, followed by XGBoost, FTTransformer, TabulaRNN, decision tree, TCN, RNN, LSTM, and SVM (Table 2). The ranking is descriptive and was not used for model selection.

Table 2. AUROC of the nine candidate models in the internal test set.

Algorithm	Accuracy	AUROC	PPV	NPV
TabulaRNN	0.90	0.87	0.83	0.90
CatBoost	0.92	0.91	0.90	0.92
XGBoost	0.92	0.90	0.79	0.93
SVM	0.86	0.72	0.32	0.89
FTTransformer	0.92	0.88	0.82	0.92
DT	0.90	0.85	0.67	0.92
RNN	0.87	0.78	0.38	0.89
LSTM	0.87	0.77	0.38	0.90
TCN	0.89	0.81	0.55	0.91

AUROC, area under the receiver operating characteristic curve; SVM, support vector machine; DT, decision tree; RNN, recurrent neural network; LSTM, long short-term memory network; TCN, temporal convolutional network.

3.5. Decision Curve Analysis

Decision curve analysis in the internal test set showed that most models yielded positive net benefit relative to the treat-none strategy across portions of the low-to-moderate threshold range, and CatBoost showed a relatively stable curve across a broad range of thresholds (Figure 4). Clinical interpretation nonetheless requires specification of the action triggered by a high-risk classification and of the relative consequences of false-positive and false-negative classifications.

Accordingly, the decision curves provide exploratory evidence of potential net benefit under the evaluated threshold assumptions and do not establish clinical utility or clinical impact.

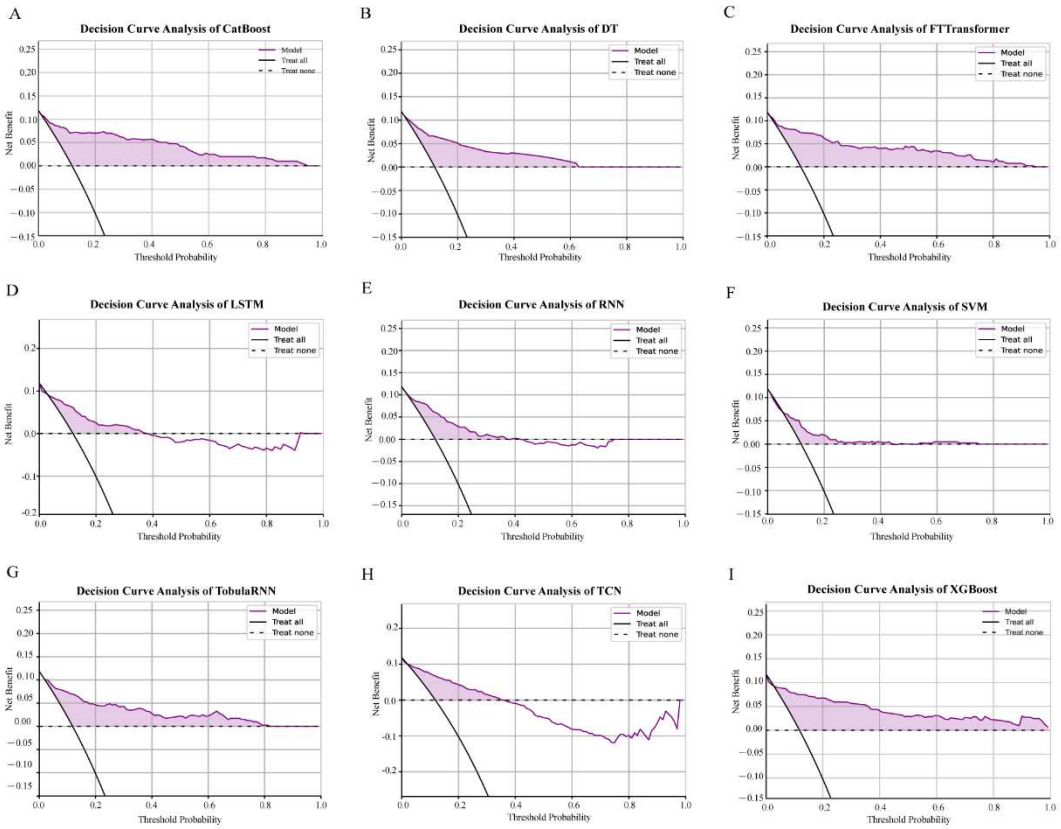


Figure 4. Decision curve analysis of the candidate prediction models in the internal test set. Decision curves show the net benefit of each model across threshold probabilities. The purple line represents the prediction model, the solid black line represents the "treat-all" strategy, and the dashed black line represents the "treat-none" strategy. Panels show decision curves for (A) CatBoost; (B) DT; (C) FTTransformer; (D) LSTM; (E) RNN; (F) SVM; (G) TabulaRNN; (H) TCN; and (I) XGBoost. DT, decision tree; LSTM, long short-term memory network; RNN, recurrent neural network; SVM, support vector machine; TCN, temporal convolutional network.

3.6. Exploratory Subgroup Comparisons of Model-Predicted Risk

Exploratory subgroup analyses compared CatBoost-predicted probabilities of 6-month eGFR-category decline across medication-use and sex groups (Figure S1). Patients using SSN had significantly higher model-predicted risk than non-users, whereas no significant differences were observed according to JSB, KLX, BLJN, or SGLT2i use or sex.

Because these comparisons were based on predicted probabilities from an observational model, and because medication indicators were included among model inputs, the findings should be interpreted only as descriptive differences in model-defined baseline risk and not as treatment effects.

3.7. Exploratory DML-Adjusted CPM–Outcome Associations

Exploratory DML yielded an adjusted risk-difference estimate of −0.08 (95% CI −0.16 to −0.01) for SSN exposure; corresponding estimates were −0.03 (95% CI −0.07 to 0.02) for KLX, 0.04 (95% CI −0.04 to 0.12) for JSB, and −0.01 (95% CI −0.07 to 0.05) for BLJN (Figure S2). Only the interval for SSN excluded zero. Given the retrospective prescription-based exposure definition, small exposed groups, and inability to verify all assumptions required for causal inference, these estimates should be interpreted as exploratory adjusted associations rather than treatment effects.

The SSN finding is therefore hypothesis-generating. Prescription records do not establish adherence or treatment duration, and residual confounding, limited treatment overlap, and instability related to small exposed groups cannot be excluded. Accordingly, no causal or treatment-selection conclusion should be drawn from the DML analysis, and confirmation in prospectively designed studies would be required.

3.8. Research-Prototype Implementation

The final CatBoost model was implemented as a mini-program–based research prototype whose interface accepts baseline demographic and laboratory variables and baseline medication indicators (Figure 5A,B). The prototype was used only to demonstrate risk calculation and non-causal scenario visualization and was not evaluated as a clinical decision-support system.

Two deidentified representative records were used solely to demonstrate the software workflow; the prototype displayed predicted risks of 92.2% and 35.3% for a record with and a record without the observed outcome, respectively (Figure 5C,D). These examples do not validate calibration, usability, safety, treatment selection, or clinical impact.

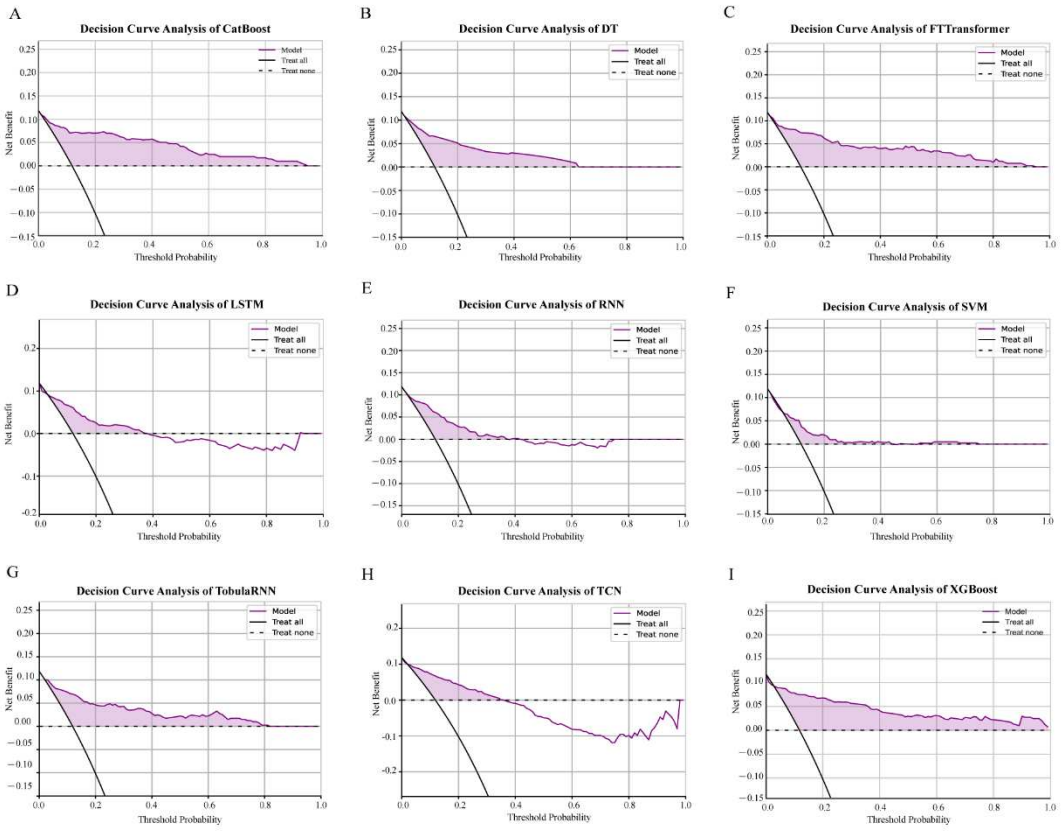


Figure 5. Mini-program research-prototype implementation of the CatBoost prediction model. (A) Baseline clinical and laboratory input interface. (B) Medication exposure input interface. (C) Example output for a deidentified representative record with the observed outcome, showing a predicted risk of 92.2%. (D) Example output for a deidentified representative record without the observed outcome, showing a predicted risk of 35.3%. These examples demonstrate interface operation only and do not constitute validation of calibration, clinical usability, treatment selection, or clinical impact. eGFR, estimated glomerular filtration rate; UACR, urinary albumin-to-creatinine ratio; SGLT2i, sodium-glucose cotransporter 2 inhibitor; SSN, Shenshuaining capsule; KLX, Keluoxin capsule; JSB, Jinshuibao capsule; BLJN, Bailing capsule.

4. Discussion

In this single-center retrospective cohort of adults with diabetes and renal dysfunction, an interpretable CatBoost model based on 15 predictors had the highest mean cross-validated AUROC in the development set and an AUROC of 0.91 in the internal test set. A post hoc top-5% resource-prioritization analysis achieved high specificity and PPV but a sensitivity of only 39.58%, indicating that this rank-based rule identifies a small very-high-risk subgroup while missing most outcome events. SHAP linked predictions to clinically familiar factors, whereas clustering, subgroup, DML, and prototype analyses were exploratory and hypothesis-generating.

A strength of this study is its focus on a pragmatic real-world population—adults with diabetes and renal dysfunction—rather than strictly adjudicated DKD. This distinction matters because definitive DKD diagnosis often requires longitudinal albuminuria and eGFR assessment, exclusion of non-diabetic kidney disease, and sometimes histological confirmation, all of which are frequently incomplete in routine electronic health records [2,3]. Use of the more conservative population definition improves transparency and reproducibility and is consistent with real-world risk-prediction studies that explicitly acknowledge the limitations of routinely collected data [3,25].

Within this context, individualized prediction remains clinically relevant. Current guidelines emphasize early identification and comprehensive management of high-risk patients based on eGFR, albuminuria, glycemic control, blood pressure, and cardiovascular–kidney–metabolic risk [2,4]. Although SGLT2i, finerenone, and glucagon-like peptide-1 receptor agonists have expanded kidney-protective treatment options [6–9,26], residual risk persists and varies substantially among patients. Models that integrate routine laboratory data, comorbidities, and medication exposure may have potential to support risk stratification and follow-up planning after further validation.

The favorable discrimination of CatBoost is consistent with the known advantages of gradient boosting for structured clinical data, including the ability to model nonlinear relationships and interactions among heterogeneous predictors; CatBoost additionally mitigates prediction shift and handles categorical features efficiently through ordered boosting and ordered target encoding [20]. In our study, AUROC increased rapidly with the addition of the top-ranked predictors and then plateaued, suggesting that a compact predictor set preserved discrimination while facilitating implementation. This observation is consistent with recent studies showing that machine learning can improve CKD progression prediction when clinically meaningful variables and transparent validation are used [13–15].

The selected predictors were clinically interpretable. eGFR and UACR are core indicators of kidney function and damage [2], whereas HbA1c, triglycerides, hsCRP, WBC count, and RBC count reflect glycemic control, metabolic disturbance, inflammation, and hematologic status, all of which are biologically relevant to CKD progression and adverse renal outcomes [27-30]. SHAP analysis further linked model predictions to these familiar factors, supporting biological plausibility and reducing the black-box character of the model, which is essential for clinical machine-learning applications [17,21].

The medication-related findings require cautious interpretation. SSN users had higher model-predicted risk in the subgroup analysis, whereas exploratory DML yielded a negative adjusted exposure–outcome estimate. This discrepancy may reflect confounding by indication and differences in measured baseline risk, although residual confounding, limited treatment overlap, exposure misclassification, model-related factors, and chance cannot be excluded. DML can reduce bias related to measured high-dimensional covariates but cannot address unmeasured confounding or guarantee causal identification. Neither analysis demonstrates an individual treatment effect, and the SSN association requires evaluation in adequately powered prospective studies or a formal target-trial emulation [23].

Implementation of the model as a mini-program research prototype demonstrates feasibility rather than clinical deployment or demonstrated clinical utility. The prototype illustrates risk calculation and non-causal scenario visualization; it has not undergone temporal or external validation, formal usability testing, safety evaluation, workflow assessment, or prospective impact evaluation. Manipulation of medication indicators must not be used to recommend starting, stopping, or switching treatment, because a prognostic model does not estimate individual causal treatment effects.

This study has several limitations. First, the internal test set was drawn from the same center and calendar period as the development data; performance in other settings and periods is therefore unknown, and temporal or external validation is required. Second, the retrospective design is vulnerable to selection bias, informative follow-up, missing data, measurement variability, unmeasured confounding, and prescription-based misclassification of medication exposure. Requiring an eligible follow-up eGFR within 150–210 days may also have selected patients with more complete follow-up. Third, the outcome was a single eGFR-category transition rather than sustained CKD progression; 40 of the 240 outcome events (16.7%) were G1-to-G2 transitions, which may partly reflect physiological variation or regression to the mean, and eGFR was used as reported by the laboratory without recalculation [18]. Fourth, the top-5% cutoff was post hoc and rank-based, with a sensitivity of only 39.58%, and should be interpreted only as an exploratory resource-prioritization analysis rather than a transportable clinical threshold. The current analysis focused primarily on discrimination and did not include a full calibration assessment or additional test-set uncertainty metrics; therefore, the reported predicted probabilities should not be interpreted as fully calibrated absolute risks. Fifth, the sequence-model benchmarks were applied to one baseline record per patient and therefore do not demonstrate the value of temporal modeling. Sixth, CPM exposure was uncommon, and the exploratory DML analysis remained susceptible to limited treatment overlap,

residual confounding, and exposure misclassification. Finally, the mini-program has not undergone temporal or external validation, formal usability testing, safety evaluation, workflow assessment, or clinical-impact evaluation.

5. Conclusions

In this single-center retrospective cohort of adults with diabetes and renal dysfunction, an interpretable CatBoost model based on 15 predictors showed good internal discrimination for 6-month eGFR-category decline. The findings support the feasibility of risk stratification using routinely available clinical variables, but the model should be regarded as a research prediction tool pending calibration assessment, temporal or external validation, and prospective evaluation. Exploratory CPM findings represent adjusted observational associations and must not be interpreted as evidence for treatment selection.

Author Contributions: Yanru Liu and Tao Yang: Conceptualization, Writing – original draft. Yi Lu and Weifu Liang: Formal analysis, Data curation, Visualization, Writing – original draft. Jingru Gong: Supervision, Project administration, Writing – review & editing. Huiping Lu: Conceptualization, Supervision, Funding acquisition, Writing – review & editing. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Construction Project of the National Pilot Zone for the Inheritance, Innovation, and Development of Traditional Chinese Medicine in Pudong New Area (Grant No. PDZY-2025-1217), the Science and Technology Program of Shanghai Pudong Hospital, Fudan University Pudong Medical Center-Clinical Research Project, Cohort Study Special Program (Grant No. YJLC202403), and the Clinical Research Project of Shanghai Pudong Hospital (Grant No. IIT117).

Institutional Review Board Statement: This study was designed as a retrospective analysis of previously collected clinical data. The study protocol and waiver of informed consent were reviewed and approved by the Ethics Committee of Shanghai Pudong Hospital, Fudan University Pudong Medical Center (Approval No. 2025-KY-036-E01). The requirement for informed consent was waived due to the retrospective nature of the study. All procedures were conducted in accordance with the ethical principles of the Declaration of Helsinki and its subsequent amendments.

Informed Consent Statement: Patient consent was waived by the Ethics Committee because of the retrospective design of the study and the use of deidentified data.

Data Availability Statement: The data supporting the findings of this study are available from the corresponding authors upon reasonable request; they are not publicly available because they contain information that could compromise patient privacy.

Acknowledgments: None.

Conflicts of Interest: The authors declare no competing interests.

References

1. GBD 2023 Chronic Kidney Disease Collaborators. Global, regional, and national burden of chronic kidney disease in adults, 1990-2023, and its attributable risk factors: a systematic analysis for the Global Burden of Disease Study 2023. Lancet 2025, 406, 2461-2482. https://doi.org/10.1016/S0140-6736(25)01853-7.
2. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2024 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. Kidney Int. 2024, 105, S117–S314. https://doi.org/10.1016/j.kint.2023.10.018.
3. McDonald, H.I.; Shaw, C.; Thomas, S.L.; Mansfield, K.E.; Tomlinson, L.A.; Nitsch, D. Methodological challenges when carrying out research on CKD and AKI using routine electronic health records. Kidney Int. 2016, 90, 943-949. https://doi.org/10.1016/j.kint.2016.04.010.
4. Rossing, P.; Caramori, M.L.; Chan, J.C.N.; Heerspink, H.J.L.; Hurst, C.; Khunti, K.; Laffel, L.M.; Mastrandrea, L.D.; McGill, J.B.; Senior, P.A. Diabetes management in chronic kidney disease: A consensus report by the American Diabetes Association (ADA) and Kidney Disease: Improving Global Outcomes (KDIGO). Kidney Int. 2022, 102, 974-989. https://doi.org/10.1016/j.kint.2022.06.013.
5. American Diabetes Association Professional Practice Committee. 11. Chronic Kidney Disease and Risk Management: Standards of Care in Diabetes–2026. Diabetes Care 2026. https://doi.org/10.2337/dc26-s011.
6. Perkovic, V.; Jardine, M.J.; Neal, B.; Bompoint, S.; Heerspink, H.J.L.; Charytan, D.M.; Edwards, R.; Agarwal, R.; Bakris, G.L.; Bull, S.B. Canagliflozin and Renal Outcomes in Type 2 Diabetes and Nephropathy. N. Engl. J. Med. 2019, 380, 2295-2306. https://doi.org/10.1056/NEJMoa1811744.
7. Heerspink, H.J.L.; Stefánsson, B.V.; Correa-Rotter, R.; Chertow, G.M.; Greene, T.; Hou, F.-F.; Lindberg, M.; McMurray, J.J.V.; Munther, P.; Katz, R. Dapagliflozin in Patients with Chronic Kidney Disease. N. Engl. J. Med. 2020, 383, 1436-1446. https://doi.org/10.1056/NEJMoa2024810.
8. The EMPA-KIDNEY Collaborative Group. Empagliflozin in Patients with Chronic Kidney Disease. N. Engl. J. Med. 2023, 388, 117-127. https://doi.org/10.1056/NEJMoa2204233.
9. Perkovic, V.; Tuttle, K.R.; Rossing, P.; Mahaffey, K.W.; Bannon, C.; de Boer, R.A.; Charytan, D.M.; Dwyer, J.P.; Garvey, W.T.; Hovingh, G.K. Effects of Semaglutide on Chronic Kidney Disease in Patients with Type 2 Diabetes. N. Engl. J. Med. 2024, 391, 109-121. https://doi.org/10.1056/NEJMoa2403347
10. Gao, Y.; Li, Z.; Wang, Y.; Zhang, H.; Huang, K.; Fu, Y.; Xu, S.; Li, Q.; Liu, X.; Zhang, G. Analysis of clinical evidence on traditional Chinese medicine for the treatment of diabetic nephropathy: a comprehensive review with evidence mapping. Front. Endocrinol. 2024, 15, 1324782. https://doi.org/10.3389/fendo.2024.1324782
11. Zhao, J.; Wang, S.; Li, X.; Zhang, G.; Xu, Y.; Zheng, X.; Guo, J.; Zhang, Z. A Prospective, Multicentered, Randomized, Double-Blind, Placebo-Controlled Clinical Trial of Keluoxin Capsules in the Treatment of Microalbuminuria in Patients with Type 2 Early Diabetic Kidney Disease. J. Integr. Complement. Med. 2024, 30, 185-195. https://doi.org/10.1089/jicm.2022.0809
12. Tao, Y.; Luo, R.; Xiang, Y.; Lei, M.; Peng, X.; Hu, Y. Use of bailing capsules (cordyceps sinensis) in the treatment of chronic kidney disease: a meta-analysis and network pharmacology. Front. Pharmacol. 2024, 15, 1342831. https://doi.org/10.3389/fphar.2024.1342831
13. Tangri, N.; Ferguson, T.W.; Bamforth, R.J.; Leon, S.J.; Arnott, C.; Mahaffey, K.W.; de Boer, I.H.; Jardine, M.J.; Perkovic, V.; Neuen, B.L. Machine learning for prediction of chronic kidney disease progression: validation of the Klinrisk model in the CANVAS program and CREDENCE trial. Diabetes Obes. Metab. 2024, 26, 3371-3380. https://doi.org/10.1111/dom.15678
14. Khalid, F.; Alsadoun, L.; Khilji, F.; Mushtaq, M.; Eze-odurukwe, A.; Mushtaq, M.M.; Ali, H.; Farman, R.O.; Ali, S.M.; Fatima, R. Predicting the Progression of Chronic Kidney Disease: A Systematic Review of Artificial Intelligence and Machine Learning Approaches. Cureus 2024, 16, e60145. https://doi.org/10.7759/cureus.60145.

15. Zheng, J.; Li, X.; Zhu, J.; Guan, S.; Zhang, S.; Wang, W. Interpretable machine learning for predicting chronic kidney disease progression risk. Digit. Health 2024, 10, 20552076231224225. https://doi.org/10.1177/20552076231224225.
16. Feuerriegel, S.; Frauen, D.; Melnychuk, V.; Schweisthal, J.; Hess, S.; Curth, A.; Bauer, S.; Kilbertus, N.; Kohane, IS.; van der Schaar, M. Causal machine learning for predicting treatment outcomes. Nat. Med. 2024, 30, 958-968. https://doi.org/10.1038/s41591-024-02902-1.
17. Collins, G.S.; Moons, K.G.M.; Dhiman, P.; Riley, R.D.; Beam, A.L.; Van Calster, B.; Liu, Z.; Verbakel, J.Y.; Van Smeden, M.; Boulesteix, A.-L.; et al. TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024, 385, e078378. https://doi.org/10.1136/bmj-2023-078378.
18. Inker, L.A.; Eneanya, N.D.; Coresh, J.; Tighiouart, H.; Wang, D.; Sang, Y.; Crew, D.C.; Doria, A.; Estrella, M.M.; Froissart, M.; et al. New Creatinine- and Cystatin C-Based Equations to Estimate GFR without Race. N. Engl. J. Med. 2021, 385, 1737-1749. https://doi.org/10.1056/NEJMoa2102953
19. Stekhoven, D.J.; Bühlmann, P. MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics 2012, 28, 112-118. https://doi.org/10.1093/bioinformatics/btr597
20. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Doroshuk, A.; Gulin, A. CatBoost: Unbiased Boosting with Categorical Features. Adv. Neural Inf. Process. Syst. 2018, 31. https://doi.org/10.48550/arXiv.1810.11363.
21. Lundberg, S.M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. Adv. Neural Inf. Process. Syst. 2017, 30. https://doi.org/10.48550/arXiv.1705.07874.
22. Chernozhukov, V.; Chetverikov, D.; Demirer, M.; Duflo, E.; Hansen, C.; Newey, W.; Robins, J. Double/Debiased Machine Learning for Treatment and Structural Parameters. Econ. J. 2018, 21, C1-C68. https://doi.org/10.1111/ectj.12097.
23. Hernán, M.A.; Robins, J.M. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am. J. Epidemiol. 2016, 183, 758-764. https://doi.org/10.1093/aje/kwv254.
24. Vickers, A.J.; Elkin, E.B. Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. Med. Decis. Making 2006, 26, 565-574. https://doi.org/10.1177/0272989X06295361.
25. Stempniewicz, N.; Vassalotti, J.A.; Cuddeback, J.K.; Mansfield, K.E.; McEwan, P.; Rees, J.; Nitsch, D.; Thomas, S.L. Chronic Kidney Disease Testing Among Primary Care Patients With Type 2 Diabetes Across 24 U.S. Health Care Organizations. Diabetes Care 2021, 44, 2000-2009. https://doi.org/10.2337/dc20-2715.
26. Bakris, G.L.; Agarwal, R.; Anker, S.D.; Pitt, B.; Ruilope, L.M.; Rossing, P.; Kolkhof, P.; Nowack, C.; Schloemer, P.; Joseph, A.; Filippatos, G. Effect of Finerenone on Chronic Kidney Disease Outcomes in Type 2 Diabetes. N. Engl. J. Med. 2020, 383, 2219-2229. https://doi.org/10.1056/NEJMoa2025845.
27. Mihai, S.; Codrici, E.; Popescu, I.D.; Enciu, A.; Albulescu, L.; Necula, L.G.; Mambet, C.; Anton, G.; Tanase, C. Inflammation-Related Mechanisms in Chronic Kidney Disease Prediction, Progression, and Outcome. J. Immunol. Res. 2018, 2018, 2180373. https://doi.org/10.1155/2018/2180373.
28. Tsai, Y.-C.; Hung, C.-C.; Kuo, M.-C.; Tsai, J.-C.; Yeh, S.-M.; Hwang, S.-J.; Chiu, Y.-L.; Kuo, H.-T.; Chang, J.-M.; Chen, H.-C. Association of hsCRP, White Blood Cell Count and Ferritin with Renal Outcome in Chronic Kidney Disease Patients. PLoS One 2012, 7, e52775. https://doi.org/10.1371/journal.pone.0052775.
29. Gluba-Brzózka, A.; Franczyk, B.; Olszewski, R.; Rysz, J. The Influence of Inflammation on Anemia in CKD Patients. Int. J. Mol. Sci. 2020, 21, 725. https://doi.org/10.3390/ijms21030725.
30. Palaka, E.; Grandy, S.; van Haalen, H.; McEwan, P.; Darlington, O. The Impact of CKD Anaemia on Patients: Incidence, Risk Factors, and Clinical Outcomes—A Systematic Literature Review. Int. J. Nephrol. 2020, 2020, 7692376. https://doi.org/10.1155/2020/7692376.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterialFinal.docx