Balkan Med J 2026;43:372-382

Reduction in Albuminuria with Dapagliflozin Exhibits Substantial Interindividual Variability: A Real-World Study Utilizing a Big Data Approach

Jinfang Song[1,2], Yongru Zhuang[1,2], Chenling Wu[3], Yi Xu[4], Ya Chen[5], Tao Wang[6], Xiaoxing Yin[2], Fen Xie[1]

[1]Department of Pharmacy, Affiliated Hospital of Jiangnan University, Jiangsu, China
[2]Jiangsu Key Laboratory of New Drug Research and Clinical Pharmacy, Xuzhou Medical University, Jiangsu, China
[3]Department of Data Information, Nanjing Jiangbei New Area biopharmaceutical Public Service Platform Co., Ltd, Jiangsu, China
[4]Phase I Clinical Trial Center, Affiliated Lianyungang Hospital of Xuzhou Medical University, Jiangsu, China
[5]Department of Endocrinology, Affiliated Hospital of Jiangnan University, Jiangsu, China
[6]Department of Pharmacy, Affiliated Hospital of Xuzhou Medical University, Jiangsu, China

Background: Differences in the albuminuria-lowering effect of dapagliflozin and the associated factors require further exploration.

Aims: To evaluate interindividual variability in the albuminuria-lowering effect of dapagliflozin and identify factors associated with the albuminuria response to dapagliflozin.

Study Design: Retrospective observational cohort study.

Methods: This real-world study, utilizing a big data approach, included patients diagnosed with type 2 diabetic kidney disease (T2DKD) who received dapagliflozin treatment for a minimum of 3 months. Using the change in urinary albumin-creatinine ratio (UACR) after 3 months of dapagliflozin treatment as the primary surrogate endpoint, patients with a UACR reduction of > 30% were categorized as responders, whereas the remaining patients were classified as non-responders. Logistic regression analysis was conducted to explore factors affecting the albuminuria response to dapagliflozin, and receiver operating characteristic (ROC) curves were plotted to evaluate the discriminative ability of the associated factors.

Results: Among 10,860 patients with T2DKD, 4,487 (41.32%) were classified as non-responders to dapagliflozin. Logistic regression analysis identified the duration of T2DM ($p < 0.001$), baseline hemoglobin A1c (HbA1c) ($p = 0.002$), baseline systolic blood pressure (SBP, $p = 0.004$), and baseline UACR ($p < 0.001$) as factors associated with the albuminuria response to dapagliflozin. The areas under the ROC curve for baseline HbA1c, baseline SBP, baseline UACR, duration of T2DM, and the combination of these four factors in assessing the efficacy of dapagliflozin in lowering albuminuria were 0.508, 0.531, 0.563, 0.603, and 0.646, respectively.

Conclusion: Significant interindividual variability was observed in the albuminuria-lowering effect of dapagliflozin in the Chinese population. After adjustment for potential confounders, the duration of T2DM, baseline HbA1c, baseline SBP, and baseline UACR were identified as factors associated with the albuminuria response to dapagliflozin. However, their discriminative ability was limited and inadequate for clinical use according to the ROC analysis.

Corresponding author: Xiaoxing Yin, Jiangsu Key Laboratory of New Drug Research and Clinical Pharmacy, Xuzhou Medical University, Jiangsu, China

e-mail: xiaoxing_yin@163.com

Corresponding author: Fen Xie, Department of Pharmacy, Affiliated Hospital of Jiangnan University, Jiangsu, China

e-mail: 9862022071@jiangnan.edu.cn

Received: December 25, 2025 **Accepted:** March 12, 2026 **Available Online Date:** July 1, 2026 • **DOI:** 10.4274/balkanmedj.galenos.2026.2025-12-261

Available at www.balkanmedicaljournal.org

ORCID iDs of the authors: J.S. 0000-0002-1138-6186; Y.Z. 0009-0001-2047-8322; C.W. 0009-0000-2525-9519; Y.X. 0009-0007-8362-2364; Y.C. 0009-0002-6554-9574; T.W. 0009-0006-5162-9863; X.Y. 0000-0001-5558-5534; F.X. 0009-0006-0615-2453.

Cite this article as: Song J, Zhuang Y, Wu C, et al. Reduction in Albuminuria with Dapagliflozin Exhibits Substantial Interindividual Variability: A Real-World Study Utilizing a Big Data Approach. Balkan Med J. 2026;43:372-382

INTRODUCTION

Diabetic kidney disease (DKD) is one of the most prevalent and severe microvascular complications of diabetes mellitus, characterized by a high incidence and an insidious onset.[1] Epidemiological data demonstrate that 30–40% of patients with type 2 diabetes mellitus (T2DM) may progress to type 2 DKD (T2DKD), potentially advancing to end-stage renal disease (ESRD) and imposing a substantial burden on healthcare systems.[2] Sodium–glucose cotransporter 2 inhibitors (SGLT2is) delay the decline in estimated glomerular filtration rate (eGFR), reduce albuminuria, and improve renal composite endpoints, exhibiting renoprotective effects independent of glucose reduction.[3] In clinical practice, the American Diabetes Association (ADA) and the Chinese Diabetes Society designate SGLT2is as a class A recommendation for the treatment of patients with T2DKD.[4,5]

Dapagliflozin was the first SGLT2i approved for clinical use, with excellent renoprotective effects supported by substantial evidence in addition to its glucose-lowering effect. The DECLARE-TIMI 58 trial found that dapagliflozin consistently improved albuminuria, delayed the decline in eGFR by 46%, and reduced the risk of the cardiorenal secondary composite outcome (eGFR decline of ≥ 40% to < 60 mL/min/1.73 m², progression to ESRD, or death from cardiorenal causes) by 24%, and the renal-specific composite outcome (eGFR decline of ≥ 40% to < 60 mL/min/1.73 m², progression to ESRD, or death from renal causes) by 47%.[6,7] The results of the DAPA-chronic kidney disease (CKD) study suggested that in patients with CKD, with and without T2DM, dapagliflozin reduced albuminuria by 29% and reduced the risk of the primary endpoints (sustained decline in eGFR ≥ 50%, progression to ESRD, or death from cardiorenal causes) by 39% and the kidney-specific composite endpoints (sustained decline in eGFR ≥ 50%, progression to ESRD, or kidney-related death) by 44%.[8] The renoprotective effects of SGLT2is, independent of glucose-lowering efficacy, may be mediated through mechanisms such as reducing oxidative stress in the kidney, inhibiting elevated angiotensinogen levels, and improving the renal inflammatory microenvironment.[9-11] However, the responsiveness of different individuals to the renoprotective effects of dapagliflozin appears to be inconsistent, potentially leading to delayed disease management, reduced treatment adherence, and inefficient utilization of healthcare resources. One study retrospectively analyzed data from several international multicenter phase III randomized controlled trials involving predominantly White and Black populations. The results indicated that approximately 46% of patients with T2DKD did not respond to treatment with dapagliflozin and that reliable predictors of responsiveness were lacking.[12] In addition, a prospective clinical trial conducted in a Norwegian population further confirmed significant interindividual differences in the albuminuria response to dapagliflozin among patients with T2DM and albuminuria, and these differences were reproducible.[13] Based on clinical experience and international research data, we hypothesized that the albuminuria-lowering effect of dapagliflozin may exhibit significant and reproducible interindividual variability. Healthcare big data represents an important application branch of big data technology in the medical field, characterized by large data volumes, diverse data types, and comprehensive medical information.[14,15] Conventional clinical trials often fail to reflect drug effects under real-world conditions because of limited sample sizes and controlled study environments. In contrast, healthcare big data incorporating comprehensive real-world evidence encompasses diverse populations and clinical contexts, providing an opportunity to evaluate interindividual variability in drug response while allowing adjustment for potential confounding factors.[16] Real-world studies utilizing big-data approaches therefore provide a more reliable method for investigating interindividual variability in drug effects.

In summary, clinical practice and international studies conducted in different populations suggest individual variability in the albuminuria-lowering effect of dapagliflozin; however, the interindividual variability and associated factors in the Chinese Han population remain to be further explored. Therefore, a real-world study utilizing a big-data approach is required to evaluate the interindividual variability in the albuminuria-lowering effect of dapagliflozin and to identify the associated factors.

MATERIALS AND METHODS

Data source

The study population consisted of patients with T2DKD who received treatment at healthcare institutions across Jiangsu Province between January 2020 and December 2023. The case data were obtained from the standardized repository of the National Healthcare Big Data (East) Center of China. The data sheets extracted mainly included medical record face sheets, outpatient and emergency department visit records, patient admission records, outpatient and emergency medication prescriptions, inpatient medication prescription records, and laboratory results. This study is a retrospective observational analysis based exclusively on the secondary use of fully anonymized clinical data. Access to and use of the data were officially authorized by the Jiangsu Commission of Health. All data were irreversibly de-identified prior to analysis, and no identifiable personal information was accessible to the investigators.

Data collection

A unique identity document (ID) was assigned to each patient by encrypting their identification card numbers to link the various data tables for each patient case, and the following information was collected: (1) Personal information: personal ID, gender, age, body mass index (BMI), duration of T2DM, family history of T2DM, smoking history, and drinking history; (2) Diagnostic information: date of diagnosis, diagnosis name, and diagnosis code; (3) Test information: test date, test item name, test value, and the corresponding unit; the indicators collected included hemoglobin A1c (HbA1c), fasting plasma glucose (FPG), systolic blood pressure (SBP), diastolic blood pressure (DBP), eGFR, and urinary albumin-creatinine ratio (UACR); (4) Medication information: medication date, drug name, dosage, and drug code.

Data cleaning

The collected patient data were cleaned prior to analysis. During the verification of personal information, duplicate records were identified by screening the uniqueness of patient IDs, and negative values and implausible data for age, BMI, and T2DM duration were removed (e.g., T2DM duration exceeding age). For diagnosis and medication data, the diagnosis date and medication date were checked for plausibility, and duplicate records and logical outliers (e.g., diagnosis dates earlier than the date of birth) were eliminated. The process of test data cleaning required the deletion of abnormal data containing non-numeric values, special characters, missing test dates, and the top 1% and bottom 1% of the values for each test indicator.

Missing data in the study were addressed using multiple imputation by chained equations to reduce bias and improve statistical efficiency. Continuous variables were imputed using predictive mean matching, and categorical variables were imputed using logistic regression. The imputation model included all primary analysis variables and auxiliary variables related to missingness. Based on the inspection of missingness patterns and clinical reasoning, the missing-at-random assumption was considered plausible, as there was no evidence that the probability of missingness was directly associated with the unobserved outcome values themselves. Twenty imputed datasets were generated to balance efficiency and variability, and the results from these datasets were combined using Rubin's rules to obtain pooled estimates. Diagnostics were performed to assess the quality of the imputations. Comparison of the distributions of observed and imputed values showed good agreement, and trace plots for iterative imputations demonstrated stable convergence across iterations. Variability between the imputed datasets was within the expected ranges, indicating that the imputed values adequately reflected the inherent uncertainty due to missing data. Sensitivity analyses comparing imputed results with complete-case analyses yielded consistent estimates, supporting the robustness of the multiple imputation approach. Overall, multiple imputation allowed proper accounting for uncertainty due to missing data, preserved the inherent variability of the dataset, and provided more reliable estimates than single-imputation methods.

Data standardization

The collected patient data were standardized. Original disease codes or names were aligned with the definitions in the International Statistical Classification of Diseases and Related Health Problems, 10th Revision. Original drug codes or names were standardized according to the unified standard codes of the China National Healthcare Security Administration Medical Insurance Service Code Standard Database. Original test names were regularized based on clinician input, and the units of test items were determined according to the most frequently occurring unit in the dataset. Test values and units were harmonized according to the corresponding unit conversion rules. Name standardization involved automatic matching followed by manual verification using the coding standardization platform of the National Healthcare Big Data (East) Center of China to ensure data consistency and accuracy.

Patient personal history (family history of T2DM, smoking history, and drinking history) was also standardized. Textual data from electronic medical records were structured into short sentences using natural language processing technology from the National Healthcare Big Data (East) Center of China. Regular expressions were then implemented in R software to extract corresponding positive and negative keywords from these sentences. Positive results (presence of family history of T2DM, smoking history, or drinking history) were assigned a value of 1, whereas negative results (absence of these histories) were assigned a value of 0. For patient gender, "man" was assigned a value of 0 and "woman" a value of 1. Using these criteria, textual information was uniformly converted into numerical data.

Data warehouse

The database generated from the cleaned and governed data served as the scientific data warehouse on which the study was based. The final dataset included personal, diagnostic, test, and medication information. Personal information comprised patient ID, characteristics, and corresponding numerical values, while the remaining sections contained patient ID, time, events, codes, or values (Figure 1).

Data extraction

To analyze interindividual variability in the albuminuria-reducing effect of dapagliflozin, the corresponding data were extracted from the data warehouse according to the following criteria: (1) The first four digits of the diagnostic code were E11.2; (2) Dapagliflozin (drug code containing XA10BKD256A001) was prescribed with ≥ 2 dosing records, with an interval of at least 3 months between the first and last records, and a dosage of 10 mg per day; (3) UACR values were available at initial dapagliflozin use (± 3 days) and after 3 months of treatment (± 3 days), with a baseline UACR ≥ 30 mg/g; and (4) age between 20 and 75 years (inclusive).

The extracted data included: (1) diagnostic information: date of diagnosis and diagnosis code; (2) personal information: personal ID, gender, age, BMI, duration of T2DM, family history of T2DM, smoking history, and drinking history; (3) test values at initial dapagliflozin use (± 3 days): HbA1c, FPG, SBP, DBP, eGFR, and UACR; (4) UACR values at 3 months of treatment (± 3 days); and (5) medication information: medication date and drug code.

Grouping

Numerous epidemiologic and clinical studies have demonstrated that reductions in UACR are associated with improved long-term renal outcomes and a decreased risk of T2DKD progression.[17] Previous studies have shown that as UACR increases, the risk of renal events in patients with diabetes increases approximately threefold, while the risk of cardiovascular events increases by about 2.5-fold, indicating a strong association between albuminuria progression and adverse cardio-renal outcomes.[18,19] A meta-analysis showed that although greater UACR reduction correlates with a better renal prognosis, a UACR reduction of more than 30% significantly lowers the risk of eGFR decline and ESRD.[20] Clinical trials of ACE inhibitors or ARBs in patients with type 2 diabetes similarly demonstrated

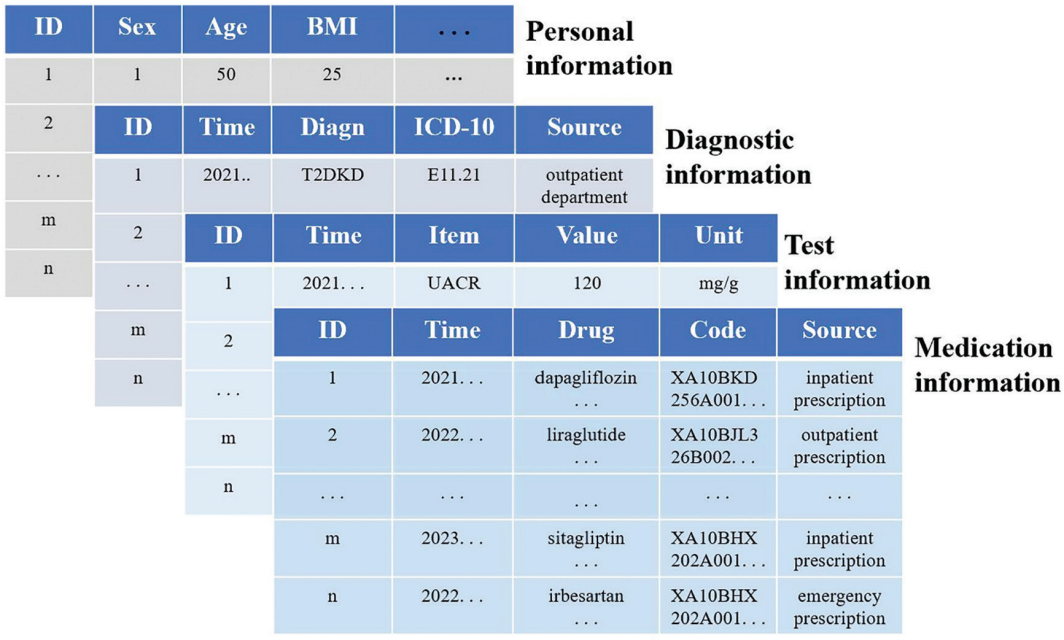

FIG. 1. Overview of the main data tables in the scientific data warehouse.

T2DKD, type 2 diabetic kidney disease; UACR, urinary albumin-creatinine ratio; BMI, body mass index; ICD-10, International Statistical Classification of Diseases and Related Health Problems, 10th Revision; ID, identity document.

that reductions in UACR of more than 30% from baseline are associated with improved renal and cardiovascular outcomes.[21] A post-hoc analysis of the EMPA-REG OUTCOME trial found that each 30% decrease in UACR over the first 12 weeks of empagliflozin treatment was significantly associated with a lower hazard of long-term renal outcomes, including a composite endpoint of sustained eGFR decline ≥ 40%, progression to renal replacement therapy, or renal death [hazard ratio (HR) = 0.83, 95% confidence interval (CI): 0.78–0.89, $p < 0.001$], indicating the relevance of early albuminuria change as a predictor of later renal protection.[22] In a real-world cohort of patients receiving SGLT2 inhibitor therapy, a > 30% decline in UACR at 3 months was associated with a lower risk of major adverse renal events (> 50% eGFR reduction or development of end-stage kidney disease) during subsequent follow-up (HR = 0.66, 95% CI: 0.48–0.89).[23] Based on this evidence, the Kidney Disease: Improving Global Outcomes and ADA guidelines recommend evaluating UACR changes after 3–6 months of SGLT2 inhibitor therapy to assess renoprotective effects, considering a 30% reduction in UACR as a reference for evaluating drug efficacy and adjusting management in T2DKD.[4,24] This threshold has also been commonly applied in previous observational studies and prospective randomized clinical trials.[12,20] Consequently, UACR was adopted as the primary surrogate endpoint in this study, with the 3-month treatment period designated as the key observational timepoint. The UACR improvement rate was defined as [(baseline UACR−UACR after treatment)/baseline UACR] × 100%, and patients with a UACR improvement of > 30% were categorized as responders, whereas patients with a UACR improvement of ≤ 30% were categorized as non-responders.

Statistical analysis

SPSS 16.0 was used for statistical analysis, and GraphPad Prism 10.1.2 was used for data visualization and plotting. The distribution of count data between groups was analyzed using the chi-square test. For continuous variables, data following a normal distribution were expressed as mean ± standard deviation, and group comparisons were performed using the independent samples t-test. Skewed data were expressed as the median with the 95% CI of the population mean, and group comparisons were conducted using non-parametric tests. Logistic regression analysis was conducted to explore factors affecting the albuminuria response to dapagliflozin, and linear regression analysis was performed to assess the linear relationship between the associated factors and the improvement rate of UACR. Receiver operating characteristic (ROC) curves were plotted to evaluate the ability of the associated factors to discriminate treatment response. Prior to regression analyses, model assumptions and diagnostics were evaluated. Multicollinearity was assessed using variance inflation factors (VIFs). For logistic regression, linearity in the logit for continuous variables was examined using the Box-Tidwell test. Model fit and influential observations were evaluated using deviance and Pearson residuals, leverage values, and Cook's distance. For linear regression, assumptions of linearity, normality, and homoscedasticity were assessed using residual and Q–Q plots, and influential observations were examined using Cook's distance. Bonferroni correction was applied to control the false positive rate in multiple comparisons, and $p < 0.05$ indicated statistical significance.

RESULTS

Change in UACR after dapagliflozin treatment

A total of 545,814 patients diagnosed with T2DKD were included in the data warehouse, and 10,860 patients were included in the final outcome analysis after excluding those with non-compliant medication records, missing test indices, or incomplete baseline characteristics (Figure 2). Analysis of UACR improvement showed an overall trend of decreasing UACR after dapagliflozin treatment compared to baseline, but the response to dapagliflozin varied significantly among individuals (Figure 3). Based on the predefined surrogate endpoint, 6,373 patients (58.68%) exhibiting a UACR improvement rate of over 30% after treatment were classified as responders, while the remaining 4,487 patients (41.32%) were classified as non-responders.

Differences in baseline characteristics between responders and non-responders

We compared baseline characteristics between responders and non-responders to identify potential confounding variables. Compared with responders, non-responders exhibited a longer duration of T2DM ($p < 0.001$), higher baseline HbA1c ($p = 0.004$), baseline SBP ($p < 0.001$), baseline DBP ($p = 0.001$), baseline eGFR ($p < 0.001$), and a higher proportion of patients using angiotensin system inhibitors ($p < 0.001$), but a lower baseline UACR ($p < 0.001$) and a lower proportion of patients using oral antidiabetic drugs and insulin concurrently ($p = 0.016$). No statistically significant differences were observed between the groups for other baseline indicators (Table 1). To further explore the impact of baseline renal function on treatment response, we performed an eGFR-stratified analysis. Supplementary Table 1 presents changes in albuminuria after dapagliflozin treatment across different baseline eGFR categories. Patients with eGFR of 60–89 mL/min/1.73 m² showed the greatest reduction in UACR after treatment (42.12% ± 34.73%), followed by those with eGFR ≥ 90 mL/min/1.73 m² (29.87% ± 18.14%). In contrast, patients with eGFR < 60 mL/min/1.73 m² exhibited the smallest improvement (16.36% ± 9.82%). Consistently, the highest response rate was observed in the 60–89 mL/min/1.73 m² group.

Exploration of factors associated with albuminuria response to dapagliflozin

To account for potential confounding factors and explore those associated with albuminuria response to dapagliflozin, logistic regression analysis was performed with the response to dapagliflozin as the dependent variable, and gender, age, duration of T2DM, family history of T2DM, smoking history, drinking history, baseline HbA1c, baseline SBP, baseline DBP, baseline UACR level, and coadministration of medication as independent variables. Model diagnostics indicated no significant multicollinearity (all VIFs < 5), satisfactory linearity in the logit for continuous variables (all $p > 0.05$), and no influential outliers based on residual analyses and Cook's distance (all Cook's D < 1), with overall model fit considered

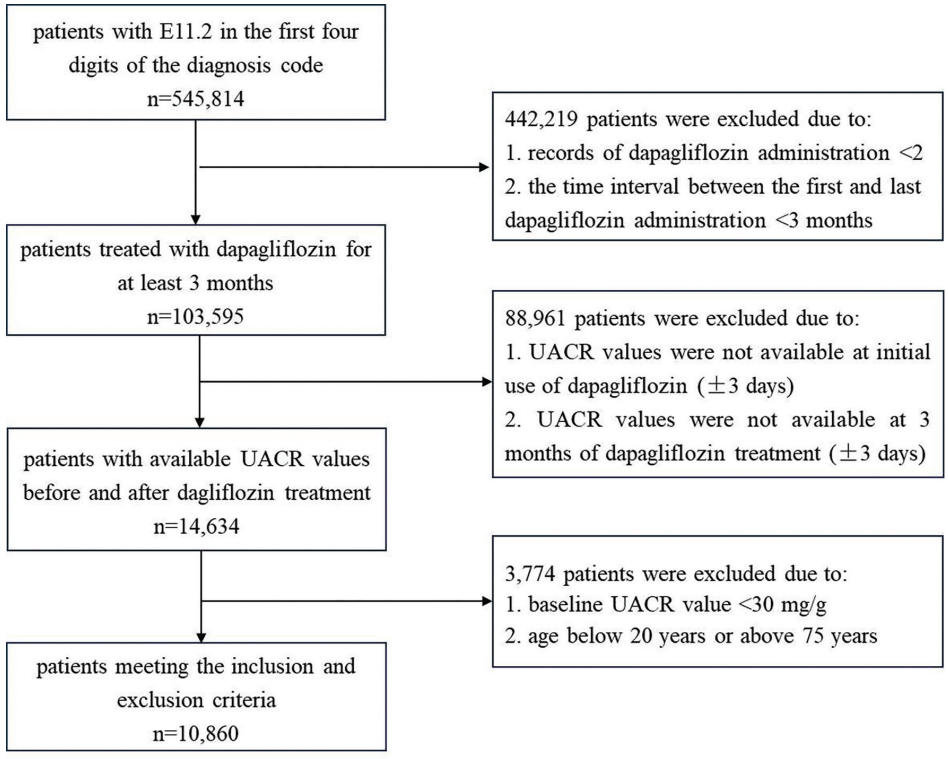

FIG. 2. Flowchart of patient inclusion and exclusion.
UACR, urinary albumin-creatinine ratio.

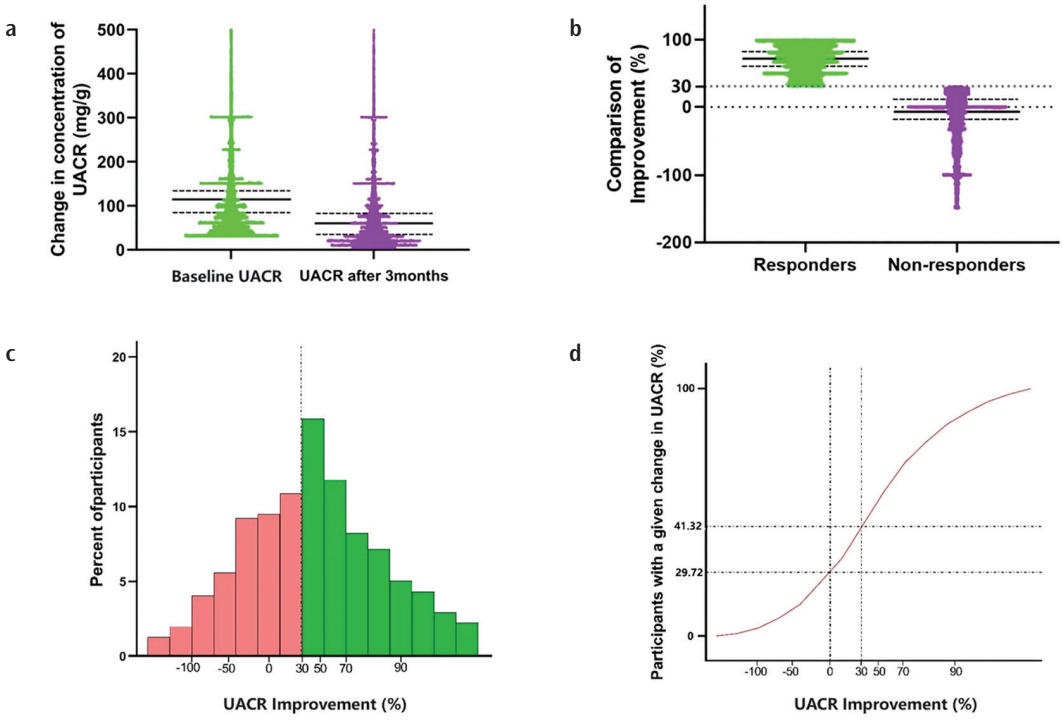

FIG. 3. Change in UACR after 3 months of dapagliflozin treatment. (a) Comparison of UACR in patients before and after dapagliflozin treatment, depicted as sample medians with 95% confidence intervals of the population median; (b) Comparison of UACR improvement rates between responders and non-responders after treatment, depicted as sample medians with 95% confidence intervals of the population median; (c) Histogram of the UACR improvement rate after treatment; (d) Cumulative distribution of UACR improvement rate after treatment. UACR, urinary albumin-creatinine ratio; the UACR improvement = [(baseline UACR−UACR after treatment)/baseline UACR] × 100%.

adequate ($p = 0.450$). The results showed that the duration of T2DM [odds ratio (OR) = 0.983, 95% CI: 0.975–0.991, $p < 0.001$], baseline HbA1c (OR = 0.964, 95% CI: 0.941–0.987, $p = 0.002$), baseline SBP (OR = 0.996, 95% CI: 0.993–0.999, $p = 0.004$), and baseline UACR (OR = 1.000, 95% CI: 1.000–1.001, $p < 0.001$) were factors associated with the albuminuria-lowering response to dapagliflozin (Table 2). Sensitivity analyses using stratified analyses by baseline UACR levels showed that the associations between clinical characteristics and response to dapagliflozin remained consistent across strata, supporting the robustness of the findings (Supplementary Table 2). Although key confounding factors were rigorously adjusted for using regression models, residual bias may still exist due to unmeasured or incompletely measured potential confounders (e.g., changes in lifestyle). However, in accordance with STROBE guidelines, retrospective study results can still provide valuable evidence for clinical practice, provided that known confounders are adequately controlled and study limitations are clearly acknowledged.[25]

In further assessing the linear relationship between the identified factors and the UACR improvement rate, linear regression analysis revealed that the duration of T2DM (β = −0.005, 95% CI: −0.007 to −0.003, $p < 0.001$), baseline HbA1c level (β = −0.009, 95% CI: −0.015 to −0.003, $p = 0.004$), baseline SBP level (β = −0.002, 95% CI: −0.003 to −0.001, $p < 0.001$), and baseline UACR level (β = 0.000, 95% CI: 0.000–0.000, $p < 0.001$) were associated with the rate of improvement in UACR after dapagliflozin treatment. However, the regression coefficients were small, indicating a weak linear relationship between these factors and UACR improvement (Supplementary Table 3, Supplementary Figure 1). Residual and Q–Q plot analyses indicated satisfactory linearity, normality, and homoscedasticity, and no influential observations were identified (all Cook's D < 1).

Ability of associated factors to discriminate treatment response

ROC curves were plotted to evaluate the ability of the associated factors to discriminate the albuminuria response to dapagliflozin. The area under the curve (AUC) for baseline HbA1c, baseline SBP, baseline UACR, and duration of T2DM in discriminating albuminuria response to dapagliflozin were 0.508 (95% CI: 0.501–0.516), 0.531 (95% CI: 0.512–0.551), 0.563 (95% CI: 0.536–0.590), and 0.603 (95% CI: 0.591–0.615), respectively. The AUC for the combined factors was 0.646 (95% CI: 0.628–0.664) (Figure 4).

TABLE 1. Comparison of Baseline Characteristics Between the Responders (n = 6.373) and Non-responders (n = 4.487).

	Groups			
Parameters	**Responders**	**Non-responders**	**SMD**	***p* value**
Sex (men/women)	3799/2243	2736/1498	0.028	0.074
Age (years)	56.62 ± 15.66	56.84 ± 15.53	0.014	0.470
BMI (kg/m^2)	25.73 ± 3.68	25.87 ± 3.74	0.038	0.152
Duration of T2DM (years)	10.70 ± 7.74	11.70 ± 8.05	0.127	< 0.001
Family history of T2DM (yes/no)	604/5769	415/4072	0.008	0.688
Smoking history (yes/no)	1264/5109	925/3562	0.019	0.318
Drinking history (yes/no)	888/5485	616/3871	0.006	0.760
HbA1c (%)	8.08 ± 2.60	8.23 ± 2.68	0.057	0.004
FPG (mmol/L)	9.72 ± 3.90	9.87 ± 4.60	0.036	0.067
SBP (mmHg)	133.50 ± 21.56	135.25 ± 23.25	0.079	< 0.001
DBP (mmHg)	81.31 ± 13.85	82.19 ± 14.73	0.062	0.001
eGFR (mL/min/1.73 m^2)	90.53 ± 31.88	93.97 ± 34.13	0.105	< 0.001
UACR (mg/g)	295.19 ± 71.89	268.88 ± 57.27	0.397	< 0.001
OADs (yes/no)	730/5643	521/3966	0.005	0.825
OADs + insulin (yes/no)	3668/2705	2477/2010	0.047	0.016
ACEI/ARB (yes/no)	4732/1641	3483/1004	0.079	< 0.001
Diuretic (yes/no)	4681/1692	3299/1188	0.002	0.950
OADs + injectable non-insulin glucose-lowering agents (yes/no)	121/6252	78/4409	0.012	0.589

BMI, body mass index; T2DM, type 2 diabetes mellitus; HbA1c, hemoglobin A1c; FPG, fasting plasma glucose; SBP, systolic blood pressure; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; UACR, urinary albumin-creatinine ratio; OADs, oral antidiabetic drugs; SMD, standardized mean difference; ACEI/ARB, angiotensin-converting enzyme inhibitors/angiotensin II receptor blockers.

TABLE 2. Logistic Regression Analysis to Identify Associated Factors.

Parameters	**OR value**	**95% CI**	***p* value**
Sex (men vs. women)	1.117	(0.973, 1.284)	0.117
Age (years)	1.004	(0.999, 1.008)	0.118
BMI (kg/m^2)	0.971	(0.902, 1.385)	0.256
Duration of T2DM (years)	0.983	(0.975, 0.991)	< 0.001
Family history of T2DM (yes/no)	0.866	(0.713, 1.053)	0.149
Smoking history (yes/no)	0.945	(0.782, 1.143)	0.562
Drinking history (yes/no)	1.107	(0.895, 1.370)	0.348
HbA1c (%)	0.964	(0.941, 0.987)	0.002
FPG (mmol/L)	0.968	(0.921, 1.076)	0.122
SBP (mmHg)	0.996	(0.993, 0.999)	0.004
DBP (mmHg)	0.996	(0.991, 1.000)	0.059
eGFR (mL/min/1.73 m^2)	0.987	(0.972, 1.002)	0.063
UACR (mg/g)	1.000	(1.000, 1.001)	< 0.001
OADs (yes/no)	2.311	(0.305, 5.530)	0.417
OADs + insulin (yes/no)	0.923	(0.609, 1.398)	0.704
ACEI/ARB (yes/no)	0.987	(0.965, 1.010)	0.078
Diuretic (yes/no)	0.940	(0.800, 1.104)	0.450
OADs + injectable non-insulin glucose-lowering agents (yes/no)	0.952	(0.836, 1.084)	0.455

BMI, body mass index; T2DM, type 2 diabetes mellitus; HbA1c, hemoglobin A1c; FPG, fasting plasma glucose; SBP, systolic blood pressure; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; UACR, urinary albumin-creatinine ratio; OADs, oral antidiabetic drugs; ACEI/ARB, angiotensin-converting enzyme inhibitors/angiotensin II receptor blockers; OR, odds ratio; CI, confidence interval.

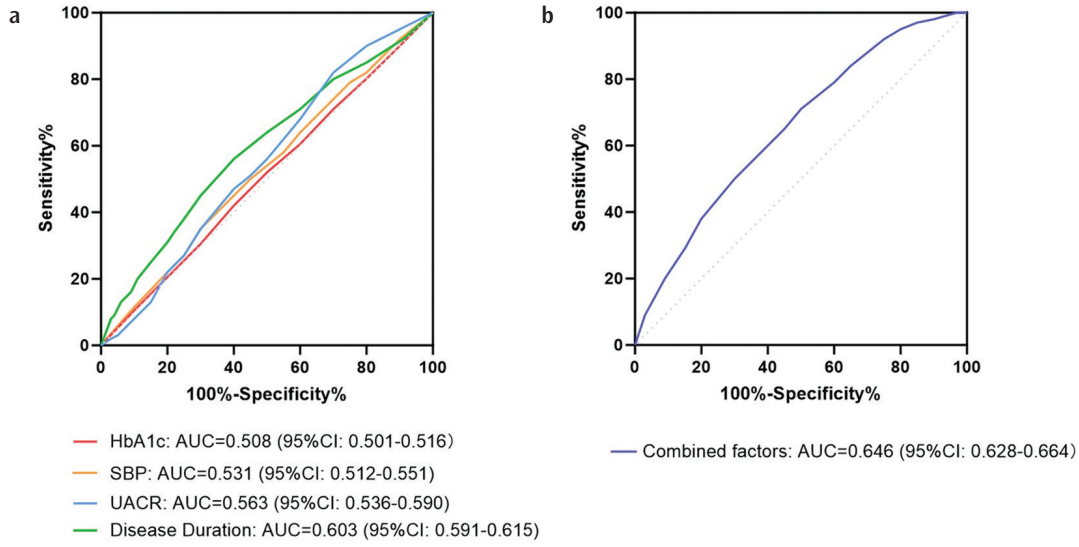

FIG. 4. ROC curves of associated factors for discriminating the albuminuria response to dapagliflozin. (a) ROC curves of duration of T2DM, UACR, SBP and HbA1c for discriminating the albuminuria response to dapagliflozin, respectively; (b) ROC curve of combined associated factors for discriminating the albuminuria response to dapagliflozin.

T2DM, type 2 diabetes mellitus; HbA1c, hemoglobin A1c; SBP, systolic blood pressure; UACR, urinary albumin-creatinine ratio; ROC, receiver operating characteristic; AUC, area under the curve.

DISCUSSION

The present real-world study, utilizing a big data approach, was conducted to evaluate interindividual variability in the albuminuria-lowering effect of dapagliflozin and to identify factors associated with the albuminuria response. A total of 10,860 patients with T2DKD treated with dapagliflozin for at least 3 months were included, and the change in UACR was considered the primary surrogate endpoint. Patients with a UACR reduction of > 30% were categorized as responders, whereas those with a reduction of ≤ 30% were classified as non-responders. The results indicated that 4,487 patients (41.32%) did not respond to dapagliflozin treatment. Baseline HbA1c, baseline SBP, baseline UACR, and duration of T2DM were associated with the albuminuria response; however, their ability to discriminate treatment response was limited. UACR is an important biomarker for assessing the progression of renal disease, with a reduction of more than 30% representing a clinically meaningful benefit. This threshold has been recommended as a criterion for evaluating the effectiveness of renoprotective drugs, such as dapagliflozin.[4,24,26] Using this criterion, previous studies reported significant individual differences in the renoprotective efficacy of dapagliflozin among Whites and Blacks, with 46% of patients with T2DKD failing to respond and a lack of reliable predictors of responsiveness.[12] A similar pattern was observed in the present study, where 41.32% of participants were non-responders after 3 months of dapagliflozin treatment. Interindividual variability may result from differences in environment, medication adherence, genetic background, and pathophysiological conditions. Based on eGFR-stratified analysis, a ≥ 30% reduction in albuminuria was more frequently achieved in patients with moderate renal function (eGFR 60–89 mL/min/1.73 m²), whereas it was less common in those with advanced renal impairment (eGFR < 60 mL/min/1.73 m²). This finding suggests that the clinical significance of a 30% UACR reduction may vary according to baseline kidney function. In patients with moderate renal impairment, a ≥ 30% reduction likely reflects meaningful improvement in glomerular hemodynamics and renal protection. In contrast, structural kidney damage in patients with advanced CKD may limit the reversibility of albuminuria, making it more difficult to achieve this threshold. In patients with preserved renal function (eGFR ≥ 90 mL/min/1.73 m²), UACR reduction after dapagliflozin treatment was smaller than in the 60–89 mL/min/1.73 m² group. This smaller relative reduction may reflect a ceiling effect, given the relatively low baseline albuminuria, which constrains the magnitude of achievable percentage decrease. Nevertheless, the absolute reduction in UACR remained clinically meaningful, indicating that patients with preserved kidney function can still benefit from SGLT2 inhibition, albeit with smaller relative changes. These findings highlight the importance of early intervention and suggest that baseline renal function should be considered when interpreting the magnitude of albuminuria reduction in clinical practice. Additionally, baseline characteristics between responders and non-responders were imbalanced. Adjusting for these factors is necessary to further explore variables associated with the albuminuria-lowering effect of dapagliflozin, supporting individualized therapy.

As this study was based on electronic medical records, the impact of unrecorded confounding factors, such as lifestyle changes, cannot be entirely ruled out. Nevertheless, methodological guidelines indicate that findings from retrospective studies remain reasonably reliable when known confounding variables are adequately controlled and limitations are explicitly acknowledged.[25] Logistic regression analysis

was performed to control for known confounders and explore factors influencing the efficacy of dapagliflozin in lowering albuminuria. Duration of T2DM, baseline HbA1c, baseline SBP, and baseline UACR were associated with the albuminuria-lowering effect. However, in studies with large sample sizes, *p* values may indicate statistical significance even with small effect sizes; therefore, effect size should be considered to interpret the practical significance of results.[27] Linear regression analysis revealed that the regression coefficients of duration of T2DM, baseline HbA1c, baseline SBP, and baseline UACR with UACR improvement were −0.005, −0.009, −0.002, and 0.000, respectively. These small effect sizes indicate a weak linear relationship. Moreover, the coefficient of determination was 0.071, suggesting that these baseline indicators accounted for only 7.1% of the variation in albuminuria response, highlighting the limited ability of these variables to explain interindividual variability in the albuminuria-lowering effect of dapagliflozin.[28]

Dapagliflozin inhibits SGLT2 in the proximal renal tubules, thereby reducing the reabsorption of sodium and glucose and promoting urinary glucose and sodium excretion.[29] This process activates the tubuloglomerular feedback mechanism, leading to constriction of the afferent arteriole, which subsequently lowers glomerular capillary pressure and alleviates the hyperfiltration state. During the early treatment period (2–4 weeks), this hemodynamic adjustment is typically accompanied by a mild and reversible decline in eGFR (generally 3–6 mL/min/1.73 m²), which stabilizes at approximately 12 weeks. The initial decline in eGFR does not indicate deterioration of renal function but rather reflects a physiological response to reduced glomerular pressure. This response is closely associated with the long-term renoprotective effects of dapagliflozin, including attenuation of eGFR decline and reduction in the risk of ESRD.[30] Reduced intraglomerular pressure decreases the permeability of the glomerular filtration membrane, thereby reducing urinary albumin leakage and improving albumin permeability, ultimately leading to improvement in UACR.[31,32]

Patients with higher baseline SBP often exhibit more severe glomerulosclerosis, tubulointerstitial fibrosis, and vascular remodeling because of prolonged hypertension. These pathological changes are difficult to reverse in the short-term and may contribute to reduced responsiveness to dapagliflozin in non-responders.[33,34] In addition, longer duration of T2DM and higher HbA1c levels are generally associated with more severe β-cell dysfunction and poorer glycemic control, leading to the accumulation of advanced glycation end products that exacerbate renal oxidative stress and inflammation.[35,36] Persistent metabolic dysregulation and inflammatory injury may reduce responsiveness to pharmacological treatment, thereby partially explaining the clinical observations in this study. The DECLARE–TIMI 58 trial reported that patients with higher baseline UACR levels showed greater reductions in urinary albumin following dapagliflozin treatment than those with lower baseline UACR levels, which is consistent with our findings.[7] However, the underlying mechanisms remain unclear.[7] One possible explanation is that some patients with T2DKD may present with relatively low baseline UACR levels because severe glomerulosclerosis and pronounced tubulointerstitial fibrosis reduce urinary protein leakage. Such patients may already have irreversible structural kidney damage, which could reduce responsiveness to renoprotective drugs such as dapagliflozin.[37]

To further evaluate the discriminatory ability of the associated factors for predicting the albuminuria response to dapagliflozin, ROC analysis was performed. The AUC values for duration of T2DM, baseline HbA1c, baseline SBP, baseline UACR, and the combination of these four factors were 0.603, 0.563, 0.531, 0.508, and 0.646, respectively. When evaluating model performance using ROC curves, an AUC value closer to 1 indicates stronger discriminative ability. Generally, an AUC of > 0.5 suggests some discriminative ability, whereas an AUC of > 0.8 indicates excellent model performance.[38] These findings suggest that although several clinical indicators are associated with the albuminuria-lowering effect of dapagliflozin, their ability to discriminate treatment response remains limited. Therefore, these results should be interpreted primarily as descriptive and hypothesis-generating rather than directly clinically actionable. Nevertheless, the findings emphasize that clinicians should remain aware of the possibility of non-response during follow-up of patients receiving dapagliflozin and highlight the need for future research aimed at individualized and precision-based therapy.

Although this study has the advantages of reliable data sources and a large sample size, several limitations should be considered. First, the data were derived from electronic medical records, making it difficult to completely exclude the influence of unrecorded factors, such as lifestyle changes and differences in medication adherence, on UACR. Second, because of the large sample size included in the study, patients with poor medication adherence could not be entirely excluded, which may have affected the accuracy of the results. Furthermore, UACR was used as a surrogate endpoint to evaluate therapeutic efficacy, with the aim of exploring changes in proteinuria and its associated factors after 3 months of dapagliflozin treatment using real-world data. Although previous studies have demonstrated that SGLT2 inhibitors can slow the decline in eGFR, the present study was unable to analyze post-treatment changes in eGFR because of limitations in data availability in large-scale real-world research. Future studies incorporating eGFR changes are needed to further elucidate the potential effects of dapagliflozin on renal function.

Although our study benefits from a large sample size and reliable electronic medical records, several additional limitations should be considered when interpreting the results. First, the study population consisted exclusively of Chinese patients with T2DKD, which may limit the generalizability of our findings to other ethnic groups or healthcare systems. Nevertheless, evidence from previously reported studies—including an international multicenter trial predominantly involving White and Black populations and a prospective trial in a Norwegian cohort—suggests that variability in the albuminuria response to dapagliflozin may be a common phenomenon, aside from potential genetic differences.[12,13] Second, during the 3-month follow-up period, variations in blood pressure, glycemic control, or initiation and dose adjustments of angiotensin-converting enzyme inhibitors/angiotensin II receptor blockers therapy may have influenced albuminuria independently of dapagliflozin treatment. For example, even modest reductions

in blood pressure or HbA1c can decrease UACR, potentially confounding the apparent drug response. Although adjustments were made for baseline values, short-term fluctuations could not be fully accounted for in this retrospective study. Moreover, real-world adherence to dapagliflozin and other therapies may vary widely. Patients with lower adherence may show smaller reductions in UACR, whereas those who improve lifestyle factors could experience larger reductions independent of the drug effect. Because electronic medical records do not fully capture these behavioral factors, some of the observed interindividual variability may reflect unmeasured confounders rather than true pharmacological differences. Finally, regression to the mean may have influenced the observed responses. Patients with higher baseline UACR are statistically more likely to exhibit larger proportional reductions, which may exaggerate treatment effects in these individuals and limit the interpretation of interindividual differences in albuminuria response. Future studies involving diverse ethnic populations and healthcare settings, preferably using prospective study designs, are needed to minimize residual confounding and validate the generalizability of these findings. Despite these limitations, the large sample size and robust data provide meaningful insights into factors associated with the albuminuria-lowering response to dapagliflozin in Chinese patients with T2DKD.

In conclusion, this real-world study using a big data approach explored interindividual variability in the albuminuria-lowering effect of dapagliflozin and the factors associated with treatment response. The results showed that 41.32% of patients with T2DKD did not respond to dapagliflozin treatment. The duration of T2DM, baseline HbA1c, baseline SBP, and baseline UACR were identified as factors associated with the albuminuria response; however, their ability to discriminate treatment response was limited. Further research is needed to identify key determinants of the albuminuria-lowering effect of dapagliflozin to support the precision use of this therapy.

Acknowledgements: We sincerely appreciate the contributions of the participants and the National Healthcare Big Data (East) Center of China.

Ethics Committee Approval: Not required.

Informed Consent: Not required.

Data Sharing Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authorship Contributions: Concept- J.S., Y.Z., X.Y., F.X.; Design- Y.Z., C.W., Y.X.; Supervision- Y.C., T.W.; Funding- J.S.; Data Collection or Processing- J.S., C.W., Y.X.; Writing- J.S., Y.Z.

Conflict of Interest: The authors declare that they have no conflict of interest.

Funding: This work was supported by the National Natural Science Foundation of China (#82204536), Top Talent Support Program for Young and Middle-aged People of Wuxi Health Committee (#HB2023064), and Jiangsu Health International Exchange Program, and Jiangsu Research Hospital Association for Precision Medication (#JY202011 and #SYHKJ-JY-2025-31).

Supplementary Tables: https://balkanmedicaljournal.org/img/files/BalkanMedJ-2025-12-261-Supplementary-Tables%282%29.pdf

Supplementary Figure: https://balkanmedicaljournal.org/img/files/BalkanMedJ-2025-12-261-Supplementary-Figure.pdf

REFERENCES

1. Fang Y, Wang B, Pang B, et al. Exploring the relations of NLR, hsCRP and MCP-1 with type 2 diabetic kidney disease: a cross-sectional study. *Sci Rep.* 2024;14:3211. [CrossRef]
2. International Diabetes Federation. Diabetes and kidney disease (2023). [CrossRef]
3. Neuen BL, Heerspink HJL, Vart P, et al. Estimated lifetime cardiovascular, kidney, and mortality benefits of combination treatment with SGLT2 inhibitors, GLP-1 receptor agonists, and nonsteroidal MRA compared with conventional care in patients with type 2 diabetes and albuminuria. *Circulation*. 2024;149:450-462. [CrossRef]
4. American Diabetes Association Professional Practice Committee. 11. Chronic kidney disease and risk management: standards of care in diabetes-2025. *Diabetes Care*. 2025;48(1 Suppl 1):239-251. [CrossRef]
5. Chinese Diabetes Society. Guideline for the prevention and treatment of diabetes mellitus in China (2024 edition). *Chin J Diabetes Mellit*. 2025;17:16-139. [CrossRef]
6. Mosenzon O, Wiviott SD, Cahn A, et al. Effects of dapagliflozin on development and progression of kidney disease in patients with type 2 diabetes: an analysis from the DECLARE-TIMI 58 randomised trial. *Lancet Diabetes Endocrinol*. 2019;7:606-617. [CrossRef]
7. Mosenzon O, Wiviott SD, Heerspink HJL, et al. The effect of dapagliflozin on albuminuria in DECLARE-TIMI 58. *Diabetes Care*. 2021;44:1805-1815. [CrossRef]
8. Wheeler DC, Stefánsson BV, Jongs N, et al. Effects of dapagliflozin on major adverse kidney and cardiovascular events in patients with diabetic and non-diabetic chronic kidney disease: a prespecified analysis from the DAPA-CKD trial. *Lancet Diabetes Endocrinol*. 2021;9:22-31. [CrossRef]
9. Karagkounis D. Cardiac and renal effects of sodium-glucose co-transporter 2 inhibitors. *Drug Ther Bull.* 2023;61:103-107. [CrossRef]
10. Woods TC, Satou R, Miyata K, et al. Canagliflozin prevents intrarenal angiotensinogen augmentation and mitigates kidney injury and hypertension in mouse model of type 2 diabetes mellitus. *Am J Nephrol*. 2019;49:331-342. [CrossRef]
11. Bailey CJ, Day C, Bellary S. Renal protection with SGLT2 inhibitors: effects in acute and chronic kidney disease. *Curr Diab Rep.* 2022;22:39-52. [CrossRef]
12. Heerspink HJL, Sjöström CD, Inzucchi SE, et al. Reduction in albuminuria with dapagliflozin cannot be predicted by baseline clinical characteristics or changes in most other risk markers. *Diabetes Obes Metab*. 2019;21:720-725. [CrossRef]
13. Petrykiv SI, Laverman GD, de Zeeuw D, Heerspink HJL. The albuminuria-lowering response to dapagliflozin is variable and reproducible among individual patients. *Diabetes Obes Metab.* 2017;19:1363-1370. [CrossRef]
14. Okada M. Big data and real-world data-based medicine in the management of hypertension. *Hypertens Res.* 2021;44:147-153. [CrossRef]
15. Piovani D, Bonovas S. Real world-big data analytics in healthcare. *Int J Environ Res Public Health*. 2022;19:11677. [CrossRef]
16. Onyejekwe ER, Sherifi D, Ching H. Perspectives on big data and big data analytics in healthcare. *Perspect Health Inf Manag*. 2024;21:1f. [CrossRef]
17. Shulman R, Yang W, Cohen DL, Reese PP, Cohen JB; CRIC study investigators. Cardiovascular and kidney outcomes of non-diabetic CKD by albuminuria severity: findings from the CRIC study. *Am J Kidney Dis.* 2024;84:742-750. [CrossRef]
18. Carrero JJ, Grams ME, Sang Y, et al. Albuminuria changes are associated with subsequent risk of end-stage renal disease and mortality. *Kidney Int.* 2017;91:244-251. [CrossRef]
19. Ninomiya T, Perkovic V, de Galan BE, et al. Albuminuria and kidney function independently predict cardiovascular and renal outcomes in diabetes. *J Am Soc Nephrol*. 2009;20:1813-1821. [CrossRef]
20. Heerspink HJL, Parving HH, Andress DL, et al. Atrasentan and renal events in patients with type 2 diabetes and chronic kidney disease (SONAR): a double-blind, randomised, placebo-controlled trial. *Lancet*. 2019;393:1936. [CrossRef]
21. Coresh J, Heerspink HJL, Sang Y, et al. Change in albuminuria and subsequent risk of end-stage kidney disease: an individual participant-level consortium meta-analysis of observational studies. *Lancet Diabetes Endocrinol*. 2019;7:115-127. [CrossRef]
22. Waijer SW, Xie D, Inzucchi SE, et al. Short-term changes in albuminuria and risk of cardiovascular and renal outcomes in type 2 diabetes mellitus: a post hoc analysis of the EMPA-REG outcome trial. *J Am Heart Assoc*. 2020;9:e016976. [CrossRef]
23. Kao YW, Yen KC, Chen SW, Chao TF, Chan YH. Early reduction in albuminuria is associated with a steeper 'dip' in initial estimated glomerular filtration rate but favourable long-term kidney outcomes in people with diabetes receiving sodium-glucose cotransporter-2 inhibitors. *Diabetes Obes Metab*. 2024;26:3868-3879. [CrossRef]

24. Kidney Disease: Improving Global Outcomes (KDIGO) Diabetes Work Group. KDIGO 2022 Clinical practice guideline for diabetes management in chronic kidney disease. *Kidney Int*. 2022;102(5 Suppl):1-127. [CrossRef]
25. Cuschieri S. The STROBE guidelines. *Saudi J Anaesth*. 2019;13(Suppl 1):31-34. [CrossRef]
26. Heerspink HJL, Greene T, Tighiouart H, et al. Change in albuminuria as a surrogate endpoint for progression of kidney disease: a meta-analysis of treatment effects in randomised clinical trials. *Lancet Diabetes Endocrinol*. 2019;7:128-139. [CrossRef]
27. Davis SL, Johnson AH, Lynch T, et al. Inclusion of effect size measures and clinical relevance in research papers. *Nurs Res*. 2021;70:222-230. [CrossRef]
28. Chicco D, Warrens MJ, Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput Sci*. 2021;7:e623. [CrossRef]
29. Matthews J, Herat L, Schlaich MP, Matthews V. The impact of SGLT2 inhibitors in the heart and kidneys regardless of diabetes status. *Int J Mol Sci*. 2023;24:14243. [CrossRef]
30. Meraz-Muñoz AY, Weinstein J, Wald R. eGFR decline after SGLT2 inhibitor initiation: the tortoise and the hare reimagined. *Kidney360*. 2021;2:1042-1047. [CrossRef]
31. Dai ZC, Chen JX, Zou R, Liang XB, Tang JX, Yao CW. Role and mechanisms of SGLT-2 inhibitors in the treatment of diabetic kidney disease. *Front Immunol*. 2023;14:1213473. [CrossRef]
32. Kanimozhi M, Bisht M, Morang S, Thapliyal S, Bassan MS, Handu S. Impact of dapagliflozin adjunctive therapy on the progression of chronic kidney disease in patients with type 2 diabetes and chronic kidney disease stages 2-5: a systematic review and meta-analysis. *Sultan Qaboos Univ Med J*. 2024;24:317-326. [CrossRef]
33. Jiang W, Wang J, Shen X, et al. Establishment and validation of a risk prediction model for early diabetic kidney disease based on a systematic review and meta-analysis of 20 cohorts. *Diabetes Care*. 2020;43:925-933. [CrossRef]
34. De Bhailis ÁM, Kalra PA. Hypertension and the kidneys. *Br J Hosp Med (Lond)*. 2022;83:1-11. [CrossRef]
35. Siddiqui K, George TP, Joy SS, Alfadda AA. Risk factors of chronic kidney disease among type 2 diabetic patients with longer duration of diabetes. *Front Endocrinol (Lausanne)*. 2022;13:1079725. [CrossRef]
36. Yan Y, Kondo N, Oniki K, et al. Predictive ability of visit-to-visit variability of HbA1c measurements for the development of diabetic kidney disease: a retrospective longitudinal observational study. *J Diabetes Res*. 2022;2022:6934188. [CrossRef]
37. Reiss AB, Jacob B, Zubair A, Srivastava A, Johnson M, De Leon J. Fibrosis in chronic kidney disease: pathophysiology and therapeutic targets. *J Clin Med*. 2024;13:1881. [CrossRef]
38. Roumeliotis S, Schurgers J, Tsalikakis DG, et al. ROC curve analysis: a useful statistic multi-tool in the research of nephrology. *Int Urol Nephrol*. 2024;56:2651-2658. [CrossRef]