

RESEARCH

Open Access



The automatic detection of diabetic kidney disease from retinal vascular parameters combined with clinical variables using artificial intelligence in type-2 diabetes patients

Shaomin Shi^{1,2†}, Ling Gao^{2†}, Juan Zhang^{1†}, Baifang Zhang³, Jing Xiao², Wan Xu², Yuan Tian^{2*}, Lihua Ni^{1*} and Xiaoyan Wu^{1,4*}

Abstract

Background Diabetic kidney disease (DKD) has become the largest cause of end-stage kidney disease. Early and accurate detection of DKD is beneficial for patients. The present detection depends on the measurement of albuminuria or the estimated glomerular filtration rate, which is invasive and not optimal; therefore, new detection tools are urgently needed. Meanwhile, a close relationship between diabetic retinopathy and DKD has been reported; thus, we aimed to develop a novel detection algorithm for DKD using artificial intelligence technology based on retinal vascular parameters combined with several easily available clinical parameters in patients with type-2 diabetes.

Methods A total of 515 consecutive patients with type-2 diabetes mellitus from Xiangyang Central Hospital were included. Patients were stratified by DKD diagnosis and split randomly into either the training set (70%, $N = 360$) or the testing set (30%, $N = 155$) (random seed = 1). Data from the training set were used to develop the machine learning algorithm (MLA), while those from the testing set were used to validate the MLA. Model performances were evaluated.

Results The MLA using the random forest classifier presented optimal performance compared with other classifiers. When validated, the accuracy, sensitivity, specificity, F1 score, and AUC for the optimal model were 84.5%(95% CI 83.3–85.7), 84.5%(82.3–86.7), 84.5%(82.7–86.3), 0.845(0.831–0.859), and 0.914(0.903–0.925), respectively.

Conclusions A new machine learning algorithm for DKD diagnosis based on fundus images and 8 easily available clinical parameters was developed, which indicated that retinal vascular changes can assist in DKD screening and detection.

Keywords Diabetic kidney disease, Diabetic retinopathy, Type 2 diabetes, Fundus photography, Artificial intelligence

[†]Shaomin Shi, Ling Gao and Juan Zhang contributed equally to this work.

*Correspondence:

Yuan Tian
tianyuanet@126.com
Lihua Ni
1047772516@qq.com
Xiaoyan Wu
wuxiaoyan2k6@whu.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Diabetic kidney disease (DKD) has already become the single largest cause of end-stage kidney disease (ESRD), and it was reported that even early stages of DKD presented an increased risk of cardiovascular disease [1]. Early detection of diabetic kidney disease will allow appropriate interventions and thus substantially reduce the health-care burden. Detection of diabetic kidney disease depends on measurement of albuminuria or estimated glomerular filtration rate (eGFR) [2], but increasing evidence has shown that some DKD patients will not develop albuminuria [3]. Moreover, the measurements of albuminuria or eGFR are not always convenient since serum or urine samples need to be obtained and they have limited precision at an earlier DKD stage [2]. Although promising new markers have been reported, none have presented better performance as a screening tool for DKD than albuminuria [4]. On the other hand, although kidney biopsy is an accurate approach to identify early disease, it is invasive and cannot be routinely used. Thus, there is a need for a noninvasive and easily available tool for diagnosing DKD.

Meanwhile, DKD and diabetic retinopathy (DR) are the most common microvascular complications of type 2-diabetes mellitus [5]. A close relationship between DR and DN has been reported in epidemiologic studies [6], since they share similar structural and physiological changes during early diabetes, which may be due to hyperglycemia, microangiopathy, inflammation, endothelial dysfunction, oxidative stress and other processes [7, 8]. As a result, DR measurements have the potential to allow optical monitoring of human microcirculation and to be a unique noninvasive predictor of DN [6]. In recent decades, a large number of studies have been conducted and have indicated that retinal changes measured from fundus photography are related to DKD, even without retinopathy [1, 9]. However, the results were inconclusive [10], with limited sensitivity, specificity, convenience and extensibility.

In the past few years, machine learning techniques have been used for the diagnosis [11, 12], prediction, and prognostic analysis of DKD [13, 14]. Most of the diagnostic models developed for DKD are based on demographic information, biochemical parameters (such as triglycerides, serum uric acid, urea nitrogen, etc.), genetic information and so on [11, 12]. However, biochemical parameters depend on the collection of blood and urine samples, so the model does not have a prominent advantage over the current diagnosis of DKD (based on urinary microalbumin). On the other hand, two studies in the past year explored the diagnostic value of fundus images for chronic kidney disease (CKD) based on artificial intelligence deep learning technology. Sabanayagam

et al. developed a diagnostic model for CKD using only fundus image information, but their diagnosis of CKD relied only on the estimated glomerular filtration rate (eGFR) [7], so the early diagnosis value is unclear. Kang Zhang et al. first used fundus photographic images to construct a diagnostic model for early DKD, and the area under the receiver operating characteristic curve (AUC) of the model was 0.800–0.864 [15]. This model directly input fundus images, but the defect lies in which specific image features are extracted; that is, the diagnostic principle is unknown.

Nowadays artificial intelligence has been increasingly used and has shown good performance in clinical diagnosis and treatment [7]. Thus, we aimed to develop a novel noninvasive detection algorithm for DKD using artificial intelligence machine learning technology based on measurable retinal vascular parameters from fundus photographs combined with several easily available clinical parameters in type-2 diabetes mellitus (T2DM) patients.

Materials and methods

Study design and population

Data from 528 sequential eligible inpatients with T2DM from Xiangyang Central Hospital from 4 January 2021 to 31 December 2021 were extracted retrospectively. Inclusion criteria were as follows: patients aged between 18–80 years old with T2DM who had UACR and eGFR data and clear retinal photographs. Full exclusion criteria were cataract or glaucoma ($n=1$), severe systemic diseases (such as end-stage renal disease, severe heart or liver disease, or malignant tumor) or acute complications of DM (such as diabetic ketoacidosis, hyperglycemic hyperosmolar state, lactic acidosis, or hypoglycemia coma) ($n=3$); and patients without gradable retinal photographs ($n=9$). Finally, 515 patients were included.

The study protocol was approved by the Ethics Committee of Xiangyang Central Hospital, an affiliated hospital of Hubei University of Arts and Science (Ethics batch number: XYSZXY-LLDD-PJ-2022–034). The study was registered at the Chinese Clinical Trial Center (Registration number: ChiCTR2200060132) and performed in accordance with the guidelines of the Declaration of Helsinki. Private personal information was removed during the process of analysis and publication. Informed consent exemptions were approved by the ethics committees.

Definition of diabetic kidney disease

T2DM was defined as fasting plasma glucose ≥ 7.0 mmol/L, 2-h postprandial plasma glucose (2hPG) ≥ 11.1 mmol/L, glycosylated hemoglobin (HbA1c) $\geq 6.5\%$, symptoms plus random plasma glucose ≥ 11.1 mmol/L [16], self-reported physician diagnosed T2DM or use of antidiabetic medications. The

urinary albumin to creatinine ratio (UACR) was calculated using spot urine collected in the morning. Albuminuria was categorized as normoalbuminuria (UACR < 30 mg/g creatinine) and albuminuria (UACR ≥ 30 mg/g creatinine). The estimated glomerular filtration rate (eGFR) was evaluated using the Chinese Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) formula as $eGFR (ml/min/1.73 m^2) = 175 \times (\text{serum creatinine } mg/dL)^{-1.154} \times (\text{Age})^{-0.203}$ ($\times 0.742$ if female) [17]. UACR and serum creatinine were measured by a Beckman Coulter AU 680 (Brea, USA). DKD was defined as an eGFR of less than 60 ml/min/1.73 m² or an albuminuria concentration of more than 30 mg/g after ruling out other possible causes of kidney injury [16]. If the data are available, only persistent albuminuria ≥ 30 mg/g, which is sustained over ≥ 90 days, will be considered “albuminuria”.

Clinical risk factors

To train models to predict DKD, 8 demographic and clinical risk factors linked to diabetic kidney disease were used as predictors, based on previous studies [2, 7, 10], and they were extracted from the hospital's electronic medical records system, including: age, gender, duration of diabetes, hypertension, history of cardiovascular and cerebrovascular disease, smoking, BMI (kg/m²), and glycosylated hemoglobin (%). Hypertension was defined as systolic blood pressure ≥ 140 mmHg or diastolic blood pressure ≥ 90 mmHg, self-reported physician-diagnosed hypertension or use of blood pressure-lowering medications. Cardiovascular and cerebrovascular diseases were defined as myocardial infarction, angina, heart failure or stroke. Body mass index (BMI) was calculated as follows: weight (kg) divided by height squared (m²). HbA1c was measured with finger capillary blood using test paper produced by Hangzhou Liwei Technology Co., Ltd., China.

Fundus photographing

Color fundus photographs were obtained using a retinal fundus camera (Canon, CR-2, Japan) at 45° of both eyes for each patient. Mydriatic agents were not applied. Macula-centered images were collected. The left eye data were used, and when unavailable, the right eye data were used. The images were obtained in JPG format. For all subjects with more than one visit, the most recent data and images were used.

The measurement was performed using a mature automatic computer program developed by the Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Xi'an Jiaotong University, China, which had been verified to be accurate and effective [1, 8, 18]. First, the deep learning algorithm was used to find the

optic disc of the fundus and separate the arteries and veins. With the optic disc as the center, blood vessels were divided into three regions: central zone (0.5–1 times the diameter of optic disc [DD]), intermediate zone (1–2 times DD), and peripheral zone (> 2 times DD). The measured parameters included nonvascular area, total vessel tortuosity, total fractal dimension and vessel caliber. For vessel caliber, the above three regions were measured and averaged, and the arteries and veins were measured separately, but ultimately, the peripheral vein caliber and peripheral arterial caliber were selected as predictors since the difference in their measurements between DKD patients and non-DKD patients was most significant.

Machine learning algorithm development

Python 3.6.13 (library, scikit-learn) was used to develop and validate the machine learning-based algorithm. Patients were stratified by DKD diagnosis and split randomly into either the training set (70%, $N=360$) or the testing set (30%, $N=155$) (random seed = 1). Data from the training set were used to develop the machine learning algorithm (MLA), while those from the testing set were used to validate the MLA. There was no overlap between the training set and the testing set. The input was measurable vascular parameters and 8 clinical parameters. The output was a binary classifier to classify the presence or absence of DKD. Four MLAs were developed using 4 different classifiers, including random forest (RF), support vector machine (SVM), gradient boosting decision tree (GBDT), and AdaBoost, to find an optimal detection model. Stratified tenfold cross-validation and grid search were used to search the optimal hyperparameters of classifiers to increase the performance of the models in the training cohort. In the RF model, Gini importance was used as a general measure of feature relevance. The accuracy, sensitivity, specificity, F1 score and area under the receiver operating characteristic curve (AUC) were calculated to evaluate the performance of the models. Predictive models were compared among classifiers, and the one with the best performance was selected (as summarized in Fig. 1).

Statistical analysis

Python 3.6.13 (library, scikit-learn) was used for development and validation of the models. STATA Version 16 (Stata Corporation, College Station, TX, USA) was used for statistical analysis. Continuous variables are presented as the mean ± standard deviation, and categorical variables are presented as percentages. Student's t test was used for continuous data, and the χ^2 test was used for categorical data. The accuracy, sensitivity, specificity

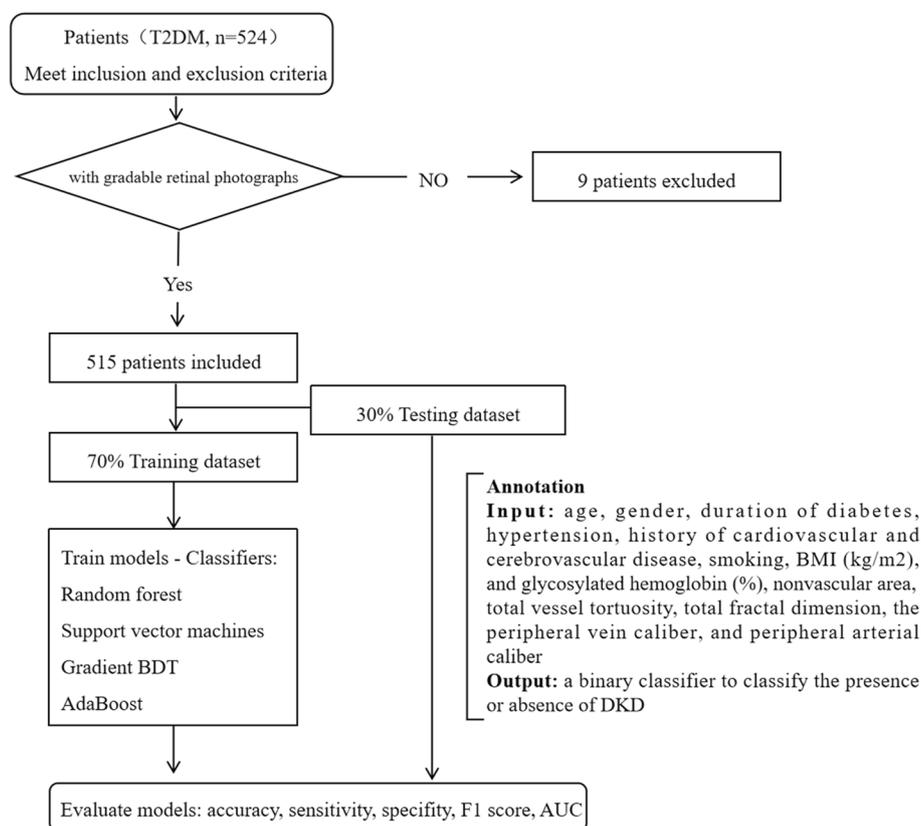


Fig. 1 The workflow for developing machine learning models to detect diabetic kidney disease in this study

F1 score, and AUC were calculated to evaluate the performance of the prediction models.

A small number of missing data (2 patients were without duration of diabetes, 17 were without BMI, and 74 were without HbA1c) were imputed and replaced by the method of k-Nearest Neighbor (k-NN) imputation. In order to correct for dataset imbalance (149 patients with DKD in the overall cohort; 28.9%), two different over-sampling algorithms were applied to all the models: random oversampling, and synthetic minority over-sampling technique (SMOTE). Overfitting bias was defined as the difference between the accuracy observed at training and accuracy at validation, and the learning curve was used to display the degree of fitting of the selected model. Statistical significance was indicated by a P value < 0.05 (two-sided).

Results

Basic characteristics and retinal measurements

The mean age of the 515 included patients was 54 years, and 329 (63.9%) were male. Moreover, 46.8% had hypertension, while 17.3% and 42.7% ever had cardiovascular/cerebrovascular diseases and smoking history,

respectively. The median UACR and eGFR were 59.6 mg/g and 110.3 ml/min/1.73 m², respectively.

Compared with T2DM patients without DKD, those with DKD were older, had a longer duration and higher BMI, and more frequently had hypertension or cardiovascular/cerebrovascular diseases and a smoking history ($P < 0.05$). Moreover, patients with DKD seemed to have a higher BMI, but the difference was not statistically significant (as shown in Table 1). Regarding the vascular parameters, patients with DKD had a higher nonvascular area and vessel tortuosity and a lower fractal dimension. Vessel calibers were increased, especially in the peripheral zone ($P < 0.05$).

Evaluations of the developed models for detecting diabetic kidney disease

After tuning of hyperparameters, a RF algorithm with SMOTE correction for data set imbalance was selected as the best model and was validated in the testing data set. A detailed description of the applied supervised learning methods is provided in Supplementary Table 1 and Supplementary Fig. 1, and the tuning of the RF model with SMOTE correction was shown in Supplementary Table 2.

Table 1 Baseline characteristics of participants

	non-DKD (n = 366)	DKD (n = 149)	Total (n = 515)	P value
Risk factors				
Age (year)	52.8 ± 0.51	57.3 ± 0.94	54.1 ± 0.46	< 0.05
Gender (Male %)	65.3	63.1	63.9	0.810
Duration (month)	71.7 ± 4.0	105.9 ± 7.0	81.6 ± 3.5	< 0.05
Hypertention	38.9	67.8	46.8	< 0.05
History (%)	13.9	26.2	17.3	< 0.05
Smoking (%)	42.5	45.0	42.7	0.799
BMI (kg/m ²)	25.5 ± 0.19	26.0 ± 0.31	25.6 ± 0.16	0.109
HbA1c (%)	8.9 ± 0.13	9.7 ± 0.19	9.1 ± 0.12	< 0.05
UACR (mg/g)	11.8 ± 0.56	176.9 ± 20.91	59.6 ± 6.90	< 0.05
eGFR(ml/min/1.73m ²)	115.2 ± 1.44	98.1 ± 3.53	110.3 ± 1.48	< 0.05
Vascular parameters				
NVArea /10 ⁴ (μm ²)	163.24 ± 0.39	172.77 ± 0.97	166.00 ± 0.44	< 0.05
Tortuosity	0.481 ± 0.0005	0.483 ± 0.0009	0.481 ± 0.0108	0.0145
FD	1.4229 ± 0.0001	1.4105 ± 0.0218	1.4193 ± 0.0001	< 0.05
VW(All) (μm)				
Central	57.10 ± 0.25	57.64 ± 0.41	57.24 ± 0.20	0.2651
Middle	52.37 ± 0.24	53.57 ± 0.39	52.72 ± 0.20	0.0076
Peripheral	42.97 ± 0.18	45.13 ± 0.32	43.60 ± 0.16	< 0.05
VW(Vein) (μm)				
Central	57.18 ± 0.25	57.64 ± 0.41	57.32 ± 0.20	0.334
Middle	52.34 ± 0.23	53.56 ± 0.40	52.68 ± 0.20	0.0066
Peripheral	42.96 ± 0.18	45.15 ± 0.31	43.60 ± 0.16	< 0.05
VW(Artery) (μm)				
Central	42.97 ± 0.18	45.13 ± 0.32	48.24 ± 0.24	< 0.05
Middle	42.97 ± 0.18	45.13 ± 0.32	47.12 ± 0.20	< 0.05
Peripheral	42.83 ± 0.19	44.40 ± 0.33	43.28 ± 0.16	< 0.05

Abbreviations: DKD Indicates diabetic kidney disease, BMI Body mass index, HbA1c Glycosylated hemoglobin, UACR Urinary albuminuria creatinine ratio, eGFR Estimated glomerular filtration rate, NVArea Non-vascular area, Tortuosity total Vessel tortuosity, FD total Fractal dimension, VW Vessel width, History History of cardiovascular and cerebrovascular disease (myocardial infarction, angina, heart failure or stroke)

It comprised 26 classification trees with a maximum number of 17 splits. One of the classification trees from this model using the RF classifier is presented in Supplementary Fig. 2. The accuracy, sensitivity, specificity, and F1 score for the optimal model in training were 90.0%(95% CI 89.5–90.5), 91.8%(91.0–92.6), 88.3%(87.7–88.9), and 0.902(0.896–0.908), respectively, and in validation were 84.5%(95% CI 83.3–85.7), 84.5%(82.3–86.7), 84.5%(82.7–86.3), and 0.845(0.831–0.859), respectively, with an AUC of 0.914(0.903–0.925) (ROC curve shown in Fig. 2). The optimal model had a good fit, with accuracies of 90.0% in training and 84.5% in validation (As shown in Supplementary Fig. 3). The importance of each variable to the optimal prediction model was analyzed by SHAP (Shapley Additive Explanations) (As shown in Fig. 3).

Supplementary analyses

We further imputed the missing data by the method of backfilling missing values and retrained the machine learning models using the random forest model with SMOTE correction for data set imbalance, and the results were similar with an overall accuracy of 82.3%(95% CI 80.8–83.8) and AUC of 0.903(0.888–0.918) (as shown in Supplementary Fig. 4a). And when the missing data were replaced by the mean value, the results remained almost consistent with those before, with an overall accuracy of 83.2%(81.5–84.9) and AUC of 0.920(0.903–0.937) (as shown in Supplementary Fig. 4b). Moreover, a few typical samples’ (including 3 DKD patients and 3 non-DKD patients) clinical indicators and fundus images were displayed in Supplementary Fig. 5.

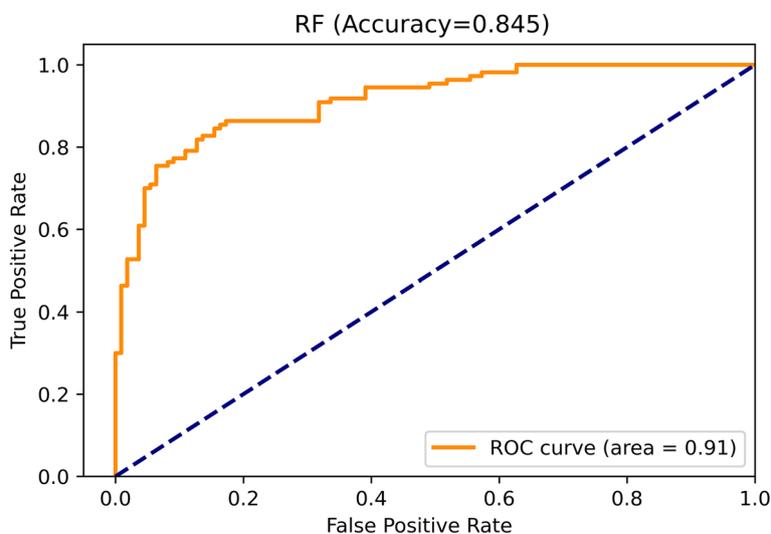


Fig. 2 ROC curve of the best model using Random Forest classifier with SMOTE correction for data set imbalance in validation

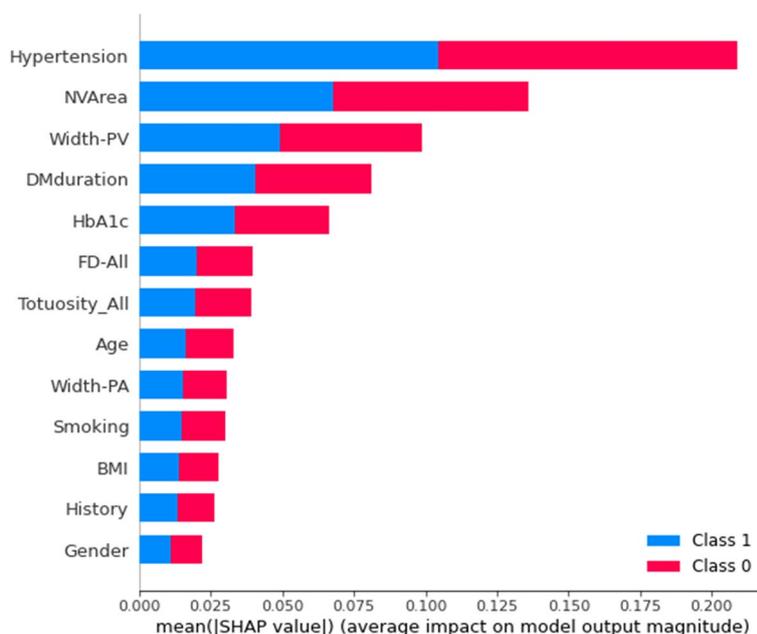


Fig. 3 Relative variable importance for the accuracy of detecting diabetic kidney disease using the random Forest classifier with SMOTE correction for data set imbalance. Abbreviations: BMI indicates body mass index; HbA1c, glycosylated hemoglobin; NVArea, non-vascular area; Tor-All, total vessel tortuosity; FD-All, total fractal dimension; Width-PA, peripheral arterial width; Width-PV, the peripheral vein width; History, history of cardiovascular and cerebrovascular disease (myocardial infarction, angina, heart failure or stroke); SMOTE, synthetic minority over-sampling technique

Discussion

DKD is the main cause of ESRD worldwide, accounting for approximately 20–40% of patients with diabetes [19, 20]. Early and accurate detection and intervention contributing to a better outcome is beneficial for the patients [13, 19]. The present detection of DKD depends

on the measurement of albuminuria or the eGFR, which are invasive and not optimal. In the present study, we developed a new detection model for DKD based on retinal vascular parameters measured from fundus images and 8 easily available clinical parameters using artificial intelligence machine learning technology.

Compared with previous studies

Early detection of DKD is of great importance to preventing the progression of nephropathy. Albuminuria or eGFR, currently used in clinical practice are not convenient enough and with limited precision at an earlier DKD stage [2]. Thus, multiple efforts have been made to explore new biomarkers for the early diagnosis of DKD. The potential biomarkers of DKD studied in recent years included kidney injury molecule 1 (KIM-1), chemokines (such as CCL19, CCL5), multi-omics related biomarkers (such as CKD273 score, various miRNAs), and so on [21]. However, all have different limitations such as poor accuracy, sensitivity, reliability and convenience, which largely influence their clinical value, and none have presented better performance as a screening tool for DKD than albuminuria [4]. Although kidney biopsy is accurate as the gold standard for the diagnosis of DKD, it is invasive and cannot be routinely used. Thus, noninvasive and easily available tools for detecting DKD are needed.

AI machine learning based on traditional risk factors has been increasingly used for the prediction of diabetic complications [22], chronic kidney disease, acute kidney injury, kidney function decline and so on, with an accuracy of 76–85.2% [13, 23–27]. For example, Makino developed a prediction model for 6-month DKD aggravation in patients with DM with 71% accuracy [20], and Belur Nagaraj et al. developed a prediction model for long-term ESRD in subjects with T2DM with 18 demographic and clinical parameters [23]. Few studies have focused on the detection of DKD. Maniruzzaman et al. developed a model with superior performance to detect DN using SVM-RBF (support vector machine-radial basis function) with 13 parameters (sex, age, BMI, DM duration, FBS, HbA1c, LDL, HDL, TGs, SBP, DBP, DM treatment, use of statins), and the accuracy and AUC were 88.7% and 0.9, respectively, when validated [22]. Satish Kumar David et al. developed a detection model for DKD with an accuracy of 93.7% using IBK and random tree classification techniques through 18 parameters (age, sex, serum albumin, sodium, potassium, urea, glucose, creatinine, HbA1c, hemoglobin, white blood cell counts, red blood cell counts, hemoglobin (%), platelet counts, SBP, DBP, hypertension, and retinopathy) [19]. However, most biochemical parameters are not convenient enough, and retinal vascular parameters have not been used in formal models, although a large number of studies have indicated that retinal changes are related to DKD [1, 9].

On the other hand, AI deep learning, which uses imaging data as a subset of machine learning, has achieved many recent advancements in the detection of various diseases over the past 20 years, including medical

kidney diseases [28]. To the best of our knowledge, only four relevant studies are available. Chin-Chi Kuo et al. developed a prediction model for CKD (GFR of <60 ml/min/1.73 m²) with 4505 ultrasound-based images of the kidney, with an accuracy of 85.6% and AUC of 0.904 without external validation [29]. Kitamura et al. developed a diagnostic algorithm for DKD based on pathological immunofluorescent images of kidneys, with an accuracy of 83.28 ± 11.64% [30]. Sabanayagam et al. developed a diagnostic model for CKD using only fundus image information, but their diagnosis of CKD relied only on the estimated glomerular filtration rate (eGFR) [7], so the early diagnosis value is unclear. Kang Zhang et al. first used fundus photographic images to construct a diagnostic model for early DKD, and the area under the receiver operating characteristic curve (AUC) of the model was 0.800–0.864 [15]. This model directly input fundus images, but the defect lies in which specific image features are extracted; that is, the diagnostic principle is unknown.

Our model was the first to use retinal vascular parameters and 8 easily available clinical parameters combined by machine learning to detect DKD [16], with an overall accuracy of 84.5%.

The strength of the present study

Although the performance of our model is not as good, it was the first to integrate data of measurable retinal vascular parameters and simple risk factor information to detect DKD, which is a noninvasive, low-cost, nonradioactive and easily available option and can be widely used as a real-time screening and diagnostic tool, since DR is suggested to be screened routinely by fundus photography annually in diabetes care [31]. Although the cost of the measurement of albuminuria or eGFR, the present detection methods of DKD, is already low, and we have not performed cost-effectiveness studies, the present model may reduce the screening cost further, since it may be integrated into the fundus cameras in the future which can screen for DKD while screening for DR at no additional cost. It has the potential to minimize unnecessary detection procedures for DKD, especially in primary care institutions with limited resources, thus improving the efficiency of the current medical system. Moreover, our model captured some information from the entire retinal images that cannot be recognized by doctors prior to clinical DKD diagnosis, such as changes in peripheral vascular caliber, vessel rarefaction and other changes; therefore, it is possible that it has some early diagnostic value beyond albuminuria or at least assists the present diagnosis method, which deserve further study. Thus, the present

model sheds light on the future development of a better detection method for DKD.

Analysis of the shortcomings and limitations

There are some limitations in this study. First, this is a retrospective study. The dataset used for the present model had a specific demographic sample, so we cannot guarantee its performance in other populations. Second, the study is cross-sectional, limiting the ability to track changes in retinal parameters or clinical variables over time, and the output of our study is binary classifier to classify the presence or absence of DKD, instead of a spectrum. Third, a sample size of 515 patients may be insufficient to capture the significant heterogeneity of DKD. Fourth, DKD may be closely associated with other factors not incorporated into the analysis, such as medication use, family history, lifestyle factors and so on. Fifth, in the present study, DKD was defined as an eGFR of less than 60 ml/min/1.73 m² or an albuminuria concentration of more than 30 mg/g and not based on pathology, and whether albuminuria is persistent (sustained over ≥ 90 days) could only be confirmed in a few patients. Sixth, the present model lacks prospective validation and external validation. Seventh, we were unable to elucidate the interpretability of the present model, which is considered as a common "black box" property of machine learning algorithms. Based on the above considerations, it is possible to optimize the model in the future. In addition, how will the present model to be fitted into the existing clinical workflows and its efficacy to the existing standard tests need to be clarified in the future.

On the other hand, the performance of the model was not good enough, which may be determined by the differences between the pathogenesis of DKD and that of DR, since they were not always consistent. Although DR and DN share similar structural and physiological changes during early diabetes [7, 8], their causes are not absolutely the same, as they have different pathogenic cytokines and different associations with neuropathy, obesity, hypertension and so on [6, 9, 32]. Thus, better studies with larger samples and studies that are based on pathological diagnoses are needed.

Conclusion

In conclusion, a new machine learning algorithm for DKD diagnosis based on fundus images and 8 easily available clinical parameters was developed, and it shed light on areas for future exploration, which indicated that fundus photography can be used as an adjunctive tool for screening and detecting DKD.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12911-023-02343-9>.

Additional file 1: Supplementary Figure 1. The ROC curves of machine learning models in validation and data imbalance correction.

Additional file 2: Supplementary Figure 2. One of the classification trees in the model using the Random Forest classifier for the detecting of DKD.

Additional file 3: Supplementary Figure 3. The learning curve of the model to detect diabetic kidney disease using Random Forest classifier with SMOTE correction.

Additional file 4: Supplementary Figure 4. ROC curve of the best model using Random Forest classifier with SMOTE correction for data set imbalance in validation. a: missing data imputed by the method of backfilling missing values; b: missing data replaced by the mean value.

Additional file 5: Supplementary Figure 5. A few typical samples' (including 3 DKD patients and 3 non-DKD patients) clinical indicators.

Additional file 6: Supplementary Table 1. The accuracy of machine learning models and data imbalance correction.

Additional file 7: Supplementary Table 2. Tuning of the machine learning model.

Additional file 8. Original data.

Additional file 9. Supplementary Python source codes. Code1: Python code for the model using RF classifier with SMOTE correction for data set imbalance. Code2: Python code for the model using SVM classifier with SMOTE correction for data set imbalance. Code3: Python code for the model using BDT classifier with SMOTE correction for data set imbalance. Code4: Python code for the model using Ada classifier with SMOTE correction for data set imbalance. Code5: Python code for the model using RF classifier with Random oversampling correction for data set imbalance. Code6: Python code for the model using SVM classifier with Random oversampling correction for data set imbalance. Code7: Python code for the model using BDT classifier with Random oversampling correction for data set imbalance. Code8: Python code for the model using Ada classifier with Random oversampling correction for data set imbalance. Code9: Python code for the model using RF classifier with no correction for data set imbalance. Code10: Python code for the model using SVM classifier with no correction for data set imbalance. Code11: Python code for the model using BDT classifier with no correction for data set imbalance. Code12: Python code for the model using Ada classifier with no correction for data set imbalance. Code13: Python code for the ROC curves of models with SMOTE correction for data set imbalance. Code14: Python code for the ROC curves of models with Random oversampling correction for data set imbalance. Code15: Python code for the ROC curves of models with no correction for data set imbalance. Code16: Python code for the model using RF classifier with SMOTE correction for data set imbalance, and imputing the missing data by the method of backfilling missing values. Code17: Python code for the model using RF classifier with SMOTE correction for data set imbalance, and imputing the missing data by means. Code18: Python code for tuning of the model using RF classifier with SMOTE correction for data set imbalance. Code19: Python code for calculating the standard deviations.

Acknowledgements

We are thankful for the Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Xi'an Jiaotong University, for measuring the retinal vascular parameters.

Authors' contributions

Shaomin Shi, Lihua Ni, and Yuan Tian designed the study, collected and analyzed the data, and wrote the first draft. Juan Zhang, Baifang Zhang, Jing Xiao, Wan Xu contributed to the data extraction and analysis. Ling Gao and Xiaoyan Wu conceived the research idea, contributed to the data interpretation, edited

the manuscript, and took the responsibility for the integrity of the data analysis. All authors approved the final version of manuscript.

Funding

None.

Availability of data and materials

The data in this study was extracted retrospectively from electronic medical record system of Xiangyang Central Hospital from 4 January 2021 to 31 December 2021. The datasets are not publicly available, but are available on request from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The study protocol was approved by the Ethics Committee of Xiangyang Central Hospital, an affiliated hospital of Hubei University of Arts and Science (Ethics batch number: XYSZYX-LLDD-PJ-2022-034). The study was registered at the Chinese Clinical Trial Center (Registration number: ChiCTR2200060132) and performed in accordance with the guidelines of the Declaration of Helsinki. Private personal information was removed during the process of analysis. The need for informed consent was waived by the Ethics Committee of Xiangyang Central Hospital, an affiliated hospital of Hubei University of Arts and Science, because of the retrospective nature of the study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Nephrology, Zhongnan Hospital of Wuhan University, 169 Donghu Road, Wuhan 430071, Hubei, China. ²Xiangyang Central Hospital, Affiliated Hospital of Hubei University of Arts and Science, Xiangyang, Hubei 441000, China. ³Department of Biochemistry, Wuhan University TaiKang Medical School (School of Basic Medical Sciences), Wuhan 430071, Hubei, China. ⁴Department of General Practice, Zhongnan Hospital of Wuhan University, 169 Donghu Road, Wuhan 430071, Hubei, China.

Received: 14 June 2023 Accepted: 16 October 2023

Published online: 30 October 2023

References

- Xu X, Sun F, Wang Q, et al. Comprehensive retinal vascular measurements: a novel association with renal function in type 2 diabetic patients in China. *Sci Rep.* 2020;10:13737.
- Khitan Z, Nath T, Santhanam P. Machine learning approach to predicting albuminuria in persons with type 2 diabetes: an analysis of the LOOK AHEAD Cohort. *J Clin Hypertens (Greenwich).* 2021;23:2137–45.
- Shi S, Ni L, Gao L, Wu X. Comparison of Nonalbuminuric and Albuminuric Diabetic Kidney Disease Among Patients With Type 2 Diabetes: A Systematic Review and Meta-Analysis. *Front Endocrinol.* 2022;13:871272.
- Hayashi Y. Detection of Lower Albuminuria Levels and Early Development of Diabetic Kidney Disease Using an Artificial Intelligence-Based Rule Extraction Approach. *Diagnostics (Basel, Switzerland).* 2019;9(4):133.
- UcugulAtilgan C, Atilgan KG, Kosekahya P, et al. Retinal microcirculation alterations in Microalbuminuric diabetic patients with and without retinopathy. *Seminars Ophthalmology.* 2021;36:406–12.
- Zhao L, Ren H, Zhang J, et al. Diabetic retinopathy, classified using the lesion-aware deep learning system, predicts diabetic end-stage renal disease in Chinese patients. *Endocrine Practice : Official J American College Endocrinology American Assoc Clinical Endocrinolog.* 2020;26:429–43.
- Sabanayagam C, Xu D, Ting DSW, et al. A deep learning algorithm to detect chronic kidney disease from retinal photographs in community-based populations. *The Lancet Digital health.* 2020;2:e295–302.
- Xu X, Gao B, Ding W, et al. Retinal image measurements and their association with chronic kidney disease in Chinese patients with type 2 diabetes: the NCD study. *Acta Diabetol.* 2021;58:363–70.
- Paterson EN, Cardwell C, MacGillivray TJ, et al. Investigation of associations between retinal microvascular parameters and albuminuria in UK Biobank: a cross-sectional case-control study. *BMC Nephrol.* 2021;22:72.
- Aronov M, Allon R, Stave D, et al. Retinal Vascular Signs as Screening and Prognostic Factors for Chronic Kidney Disease: A Systematic Review and Meta-Analysis of Current Evidence. *J Pers Med.* 2021;11(7):665.
- Huang GM, Huang KY, Lee TY, Weng J. An interpretable rule-based diagnostic classification of diabetic nephropathy among type 2 diabetes patients. *BMC Bioinformatics.* 2015;16(1):S5.
- Cho BH, Yu H, Kim KW, Kim TH, Kim IY, Kim SI. Application of irregular and unbalanced data to predict diabetic nephropathy using visualization and feature selection methods. *Artif Intell Med.* 2008;42:37–53.
- Allen A, Iqbal Z, Green-Saxena A, et al. Prediction of diabetic kidney disease with machine learning algorithms, upon the initial diagnosis of type 2 diabetes mellitus. *BMJ Open Diabetes Res Care.* 2022;10(1):e002560.
- Leung RK, Wang Y, Ma RC, et al. Using a multi-staged strategy based on machine learning and mathematical modeling to predict genotype-phenotype risk patterns in diabetic kidney disease: a prospective case-control cohort analysis. *BMC Nephrol.* 2013;14:162.
- Zhang K, Liu X, Xu J, et al. Deep-learning models for the detection and incidence prediction of chronic kidney disease and type 2 diabetes from retinal fundus images. *Nature Biomedical Engineer.* 2021;5:533–45.
- Cosentino F, Grant PJ, Aboyans V, et al. 2019 ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD. *Eur Heart J.* 2020;41:255–323.
- Changjie G, Xusheng Z, Feng H, Shuguang Q, Jianwen L, Junzhou F. Evaluation of glomerular filtration rate by different equations in Chinese elderly with chronic kidney disease. *Int Urol Nephrol.* 2017;49:133–41.
- Wang Q, Yang A, Sun F, Zhang M, Xu X, Gao B. Correlation between retinal vascular parameters and cystatin C in patients with type 2 diabetes. *Acta Diabetol.* 2021;58:1395–401.
- David SK, Rafullah M, Siddiqui K. Comparison of Different Machine Learning Techniques to Predict Diabetic Kidney Disease. *J Healthcare Engineer.* 2022;2022:7378307.
- Makino M, Yoshimoto R, Ono M, et al. Artificial intelligence predicts the progression of diabetic kidney disease using big data machine learning. *Sci Rep.* 2019;9:11862.
- Xie Z, Xiao X. Novel biomarkers and therapeutic approaches for diabetic retinopathy and nephropathy: recent progress and future perspectives. *Front Endocrinol.* 2022;13:1065856.
- Maniruzzaman M, Islam MM, Rahman MJ, Hasan MAM, Shin J. Risk prediction of diabetic nephropathy using machine learning techniques: a pilot study with secondary data. *Diabetes & metabolic syndrome.* 2021;15:102263.
- Belur Nagaraj S, Pena MJ, Ju W, Heerspink HL. Machine-learning-based early prediction of end-stage renal disease in patients with diabetic kidney disease using clinical trials data. *Diabetes Obes Metab.* 2020;22:2479–86.
- Chan L, Nadkarni GN, Fleming F, et al. Derivation and validation of a machine learning risk score using biomarker and electronic patient data to predict progression of diabetic kidney disease. *Diabetologia.* 2021;64:1504–15.
- Dong Z, Wang Q, Ke Y, et al. Prediction of 3-year risk of diabetic kidney disease using machine learning based on electronic medical records. *J Transl Med.* 2022;20:143.
- Rodriguez-Romero V, Bergstrom RF, Decker BS, Lahu G, Vakilynejad M, Bies RR. Prediction of Nephropathy in Type 2 Diabetes: an Analysis of the ACCORD Trial Applying Machine Learning Techniques. *Clin Transl Sci.* 2019;12:519–28.
- Schena FP, Anelli VW, Abbrescia DI, Di Noia T. Prediction of chronic kidney disease and its progression by artificial intelligence algorithms. *J Nephrol.* 2022;35(8):1953–71.

28. Hacking S, Bijol V. Deep learning for the classification of medical kidney disease: a pilot study for electron microscopy. *Ultrastruct Pathol.* 2021;45:118–27.
29. Kuo CC, Chang CM, Liu KT, et al. Automation of the kidney function prediction and classification through ultrasound-based kidney imaging using deep learning. *NPJ digital medicine.* 2019;2:29.
30. Kitamura S, Takahashi K, Sang Y, Fukushima K, Tsuji K, Wada J. Deep Learning Could Diagnose Diabetic Nephropathy with Renal Pathological Immunofluorescent Images. *Diagnostics (Basel, Switzerland).* 2020;10(7):466.
31. Li N, Ma M, Lai M, et al. A stratified analysis of a deep learning algorithm in the diagnosis of diabetic retinopathy in a real-world study. *J Diabetes.* 2022;14:111–20.
32. Wu Z, Yu S, Kang X, et al. Association of visceral adiposity index with incident nephropathy and retinopathy: a cohort study in the diabetic population. *Cardiovasc Diabetol.* 2022;21:32.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

