ABSTRACT

Objectives This study aims to independently and externally validate the Risk Prediction Model for Diabetic Kidney Disease (RPM-DKD) in patients with type 2 diabetes mellitus (T2DM).

Design This is a retrospective cohort study.

Setting Outpatient clinics at Lee’s United Clinics, Taiwan, China.

Participants A total of 2504 patients (average age 55.44 years, SD, 7.49 years) and 4455 patients (average age 57.88 years, SD, 8.80 years) were included for analysis in the DKD prediction and progression prediction cohorts, respectively.

Exposure The predicted risk for DKD and DKD progression for each patient were all calculated using the RPM-DKD.

Primary and secondary outcome measures The primary outcome measure was overall incidence of DKD. Secondary outcomes included DKD progression.

Results The DKD prediction cohort and progression prediction cohort consisted of patients with 2504 and 4455 T2DM, respectively. The RPM-DKD examined in this study showed moderately discriminative ability with area under the curve ranged from 0.636 to 0.681 for the occurrence of DKD and 0.620 to 0.654 for the progression of DKD. The Hosmer-Lemeshow χ² test indicated the RPM-DKD was not well calibrated for predicting the occurrence of DKD and overestimated the progression of DKD. The precision for predicting the occurrence and progression of DKD were 43.2% and 42.2%, respectively.

Conclusions On external validation, the RPM-DKD cannot accurately predict the risk of DKD occurrence and progression in patients with T2DM.

INTRODUCTION

Diabetic kidney disease (DKD) is one of the main microvascular complications of long-standing, uncontrolled type 2 diabetes mellitus (T2DM) and a main cause of preventable chronic and end-stage kidney disease worldwide.1 Paradoxically, improvements in cardiovascular survival in patients with T2DM have contributed to prolonged patient survival, which in turn lengthens time at risk for developing renal impairment.2 In China, about 20%−40% of individuals with T2DM have DKD.3 Further, progression of DKD to ESRD requiring renal replacement therapy and/or renal transplant brings economic burden and is associated with additional comorbid burden.4−6 In light of these factors, the early intervention and study of a relevant risk prediction model (RPM) for early DKD are of great clinical and societal relevance.

Clinical RPMs aim to estimate an individual’s risk of an event based on relevant contributing information.7 Currently, many prediction models have been developed to assess risk of incident diabetes, but few have been validated in subsequent analyses and applied to clinical practice.8−10 For example, one T2DM risk score, the FINDRISC, is well
known in Latin America and the Caribbean, despite limited none external validation of the model.11 An RPM should not enter clinical practice unless it has been independently and externally validated and proven to perform a useful role.12

The ability to accurately predict risk of DKD would allow for earlier recognition, and perhaps intervention, in patients with long-standing T2DM. Recently, an RPM for early DKD (RPM-DKD) was developed based on systematic review and meta-analysis of individual participant data from 20 cohorts of predominantly white populations.13 However, validation was limited in scope with a relatively small study size (n=380) and insufficient median follow-up time (t=2.9 years). The purpose of the current study was to independently and externally validate performance of the RPM-DKD in predicting the risk of incidence of DKD in patients with T2DM. In addition, although PRM-DKD was only used to predict the occurrence of DKD, we believe that it can predict the progress of DKD to a certain extent, with reason that the influencing factors of DKD progression and occurrence are similar. Therefore, the secondary purpose of this study is to evaluate the performance of this model in predicting the risk of DKD progression in patients with T2DM.

METHODS

Data sources and participants

We used outpatient data from December 2006 to October 2019 from Lee’s United Clinics (LUC) in Taiwan. LUC is a large ambulatory system, composed of six clinics providing multidisciplinary care for patients with diabetes. The Taiwan Health Insurance Plan supports four annual follow-up visits along with access to medications, diabetes supplies, diabetes self-management education, clinician visitation and primary/secondary prevention screening to patients living with diabetes. This setting provided an opportune source of robust, longitudinal data in which to validate the RPM-DKD.

Inclusion criteria for predicting the occurrence of DKD aligned with those established by the RPM-DKD13 and included: (1) patients aged 39–75 years and (2) patients without albuminuria (urinary albumin-to-creatinine ratio (UACR) <30 mg/g or albumin excretion rate (AER) <30 mg/g) and estimated glomerular filtration rate (eGFR) ≥60 mL/min/1.73 m² at baseline.

Inclusion criteria for predicting progression of DKD were patients aged 39–75 years with S1–S3 at baseline. (The criteria for S1–S3 stage presented in diagnosis criteria part).

We excluded (1) patients with less than 3 years of longitudinal follow-up, (2) those with history of acute kidney injury, primary glomerulonephritis, urinary tract infection, urinary calculi, etc, (3) patients with missing endpoints and lost to follow-up and (4) patients with end-stage renal disease (very high DKD risk) at baseline.

Diagnosis criteria

Diagnosis criteria of DKD were also consistent with foundational RPM-DKD modelling and included: (1) eGFR <60 mL/min/1.73 m² and/or (2) UACR ≥30 mg/g (or AER ≥30 mg/g) (3) present for ≥3 months caused by diabetes.14

Diagnostic criteria to clinical progression of DKD were15: (1) Patients were in a non-progression group if they maintained the same DKD stage or their condition had improved to an earlier DKD stage category and (2) Patients were in a progression group if the DKD stage category had progressed.

The stage of DKD was classified using a combination of eGFR and ACR into four stage categories16 that is: (1) low DKD risk, eGFR ≥60 mL/min/1.73 m² and ACR <30 mg/g; (2) moderate CKD risk, eGFR between 45 and 59 mL/min/1.73 m² and ACR <30 mg/g or eGFR between 60 mL/min/1.73 m² and ACR between 30 and 300 mg/g; (3) high DKD risk, eGFR between 30 and 44 mL/min/1.73 m² and ACR <30 mg/g or eGFR between 45 and 59 mL/min/1.73 m² and ACR between 30 and 300 mg/g or eGFR ≥60 mL/min/1.73 m² and ACR >300 mg/g; (4) very high DKD risk, eGFR ≤29 mL/min/1.73 m² and ACR <30 mg/g or eGFR ≤44 mL/min/1.73 m² and ACR between 30 and 300 mg/g or eGFR ≤59 mL/min/1.73 m² and ACR >300 mg/g.

Risk score calculations

The risk score model was established by Jiang et al13 and all risk factors included in the DKD risk score model were derived from a systematic review and meta-analysis of 14 prospective and 6 retrospective cohorts. The predicted risk score for each study participant was calculated using their baseline data. The baseline variables used for the risk scores were in accordance with the model: (1) age (years) divided into three categories, 39–49 scores 0, 50–59 scores 3.0 and 60–75 scores 6.0; (2) body mass index (BMI), which was calculated as the patient’s weight divided by the square of their height (kg/m²), divided into three categories, <19 scores 0, ≥19 and <25 scores 1.5, ≥25 and <30 scores 3.0; (3) smoker (defined as having smoked more than 100 cigarettes in their lifetime), non-smoker scores 0 and smoker scores 4.0; (4) diabetic retinopathy, 0 if no and 3.0 if yes; (5) haemoglobin A1c (HbA1c) divided into four categories, <7.0% (<53 mmol/mol) scores 0, 7.0%–7.9% (53–63 mmol/mol) scores 1.5, 8.0%–8.9% (64–74 mmol/mol) scores 3.0 and ≥9.0% (≥75 mmol/mol) scores 4.5; (6) systolic blood pressure (SBP) divided into four categories, <130 mm Hg scores 0, 130–139 mm Hg scores 2.0, 140–149 mm Hg scores 4.0, ≥150 mm Hg scores 6.0; (7) serum high-density lipoprotein-cholesterol (HDL-C) divided into two categories, ≥1.30 mmol/L scores 0 and <1.30 mmol/L scores 2.5; (8) triglycerides divided into two categories, ≥1.70 mmol/L scores 0 and ≥2.0 mmol/L scores 4.0; (9) UACR divided into three categories, <10 mg/g scores 0, 10.00–19.99 mg/g scores 2.0, 20.00–29.99 mg/g scores 4.0. In addition, considering that we want to predict the progression of
DKD, we continue to increase the category of UACR, that is, every increase of UACR 10 mg/g, the score increases by two points. For example, UACR between 30.00 and 39.99 mg/g scores 6.0. The coefficients in the model are shown in online supplemental appendix 1.

Data to inform score calculation were retrieved from the LUC electronic medical record. Four risk categories include: (1) relatively low (score <12.0); (2) moderate (score 12.0–15.5); (3) high (score 16.0–26.5) and (4) very high (score >27.0).

**Sample size and missing data**

Following the simulation-based sample size calculations for external validation of clinical prediction models, the anticipated precisions of performance measures were estimated based on the available number of outcome events in the external validation datasets. The number of outcome events in our DKD prediction cohort was 817. When using a fixed base probability of 0.4, the minimum sample size used in this study should be 2043 cases.

Our validation cohort had missing information on age (1.3%), BMI (2.6%) and HbA1c (5.6%). The rest variables included in RPM-DKD model were complete. We used multiple imputation (10 imputations) to replace missing values by using a chained equation approach based on all candidate predictors and outcomes.

**Statistical analysis**

Descriptive statistics were generated for all variables, which were stratified by the occurrence and progression of DKD. Normally distributed continuous variables were presented as means±SD, and analysis of variance was used to assess intergroup comparisons. Medians (IQR) were used for continuous variables that were not normally distributed, and the comparison between groups were performed by Kruskal-Wallis H test. Categorical variables were represented as number of cases (N), and the intergroup rate (%) was compared with χ² test.

The clinical performance of the DKD RPM was assessed by means of discrimination and calibration. Model discrimination describes a model’s performance in distinguishing between individuals who experience an event and those who do not. Model discrimination was assessed by plotting a receiver operating characteristic (ROC) curve and calculating the area under the curve (AUC). An AUC statistic value >0.75 was regarded to represent good discrimination. Calibration assessment of an RPM describes how well predictions match observed values.

---

**Figure 1**  Flow chart of patient selection. AER, albumin excretion rate; DKD, diabetes kidney disease; eGFR, estimated glomerular filtration rate; T2DM, type 2 diabetes mellitus; UACR, urinary albumin-to-creatinine ratio.
outcomes. The calibration of the risk score predictions was assessed by plotting observed versus predicted number of patients and by calculating the Hosmer-Lemeshow $\chi^2$ statistic. Groups for observed DKD events were based on deciles for the predicted probabilities. Performance was evaluated as sensitivity, specificity and precision.

All results were presented with 95% CI. Any two-tailed p values <0.05 were considered statistically significant. All statistical analyses were performed using SPSS software, V.22 (IBM).

**Patient and public involvement**
There was no patient or public involvement in the design and conduct of the study.

**RESULTS**
In the DKD prediction cohort, a total of 2504 patients (average age 55.44 years, SD, 7.49 years), and 4455 patients (average age 57.88 years, SD, 8.80 years) were included for analysis in the DKD progression prediction cohort (figure 1). The average length of follow-up was 7.37 years (SD, 3.22 years) in the DKD prediction cohort and a total of 817 (32.6%) people had DKD during the follow-up period. The mean follow-up time in the progression prediction cohort was 7.72 years (SD, 3.10 years), and the overall progression of events in this cohort was 1563 (35.1%).

**DKD prediction cohort**

**Baseline characteristics**
The DKD prediction cohort had an average BMI of 26.1±4.0 kg/m$^2$ and 54.8% were men. The proportion of smokers and drinkers were 30.7% and 25.4%, respectively. There were significant differences in the level of education (p=0.0021) and marital status (p<0.001) between patients who eventually developed DKD and those who did not. The patients who did not develop of DKD had higher rates of secondary and college-level education and lower rates of spousal loss than the patients with DKD. Furthermore, patients who eventually developed DKD had a longer diabetes duration (median 5 years (2–9 years)) and a higher level of HbA1c (median 8.4% (7.10%–10.10%)) than those who did not. All patients showed normal albuminuria and renal function at baseline with median UACR of 8.64 mg/g (4.50–13.86 mg/g) and median eGFR of 85.50 mL/min/1.73 m$^2$ (73.08–96.13 mL/min/1.73 m$^2$). The median SBP and diastolic blood pressure (DBP) were 129 mm Hg (119–141 mm Hg) and 77 mm Hg (70–84 mm Hg), respectively. Baseline characteristics of the DKD prediction cohort are displayed in table 1.

**External validation results for DKD prediction cohort**
Of the 2504 patients, 678 (27.1%), 639 (25.5%), 1114 (44.5%) and 73 (2.9%) had risk categories of relatively low, moderate, high and very high at baseline. At the end of observation, 129 (19.0%), 175 (27.4%), 465 (41.7%), 48 (65.8%) patients in the relatively low, moderate, high and very high groups developed DKD, respectively (figure 2).

**Discrimination of the model**
According to our external validation, 16 was selected as the optimal cut-off risk score value, at which the sum of sensitivity and specificity was maximal (Youden’s index), which corresponded with Jiang et al. With a risk cut-off value of 16 points, the sensitivity would be 53.0% (95% CI 48.9% to 57.0%), specificity would be 65.7% (95% CI 63.0% to 68.3%). ROC curve of our external validation showed the area under the DKD risk score curve was 0.659 (95% CI 0.636% to 0.681%).

** Calibration of the model**
The risk scoring model was not well calibrated for predicting the occurrence of DKD, with a Hosmer-Lemeshow $\chi^2$ statistic of 16.731 (p=0.033). The calibration plot in figure 3 shows that the comparison between observed and predicted the occurrence of DKD, indicting the overestimation of risk occurred in 1–10.

**Precision of the model in predicting DKD**
The precision of the model in predicting DKD refers to the ratio of the actual number of patients with DKD to the number of patients predicted to develop DKD. According to our data, patients in the DKD prediction cohort were classified as high risk if their DKD risk score ≥16 points, with precision of 43.2% (513/1187).

**Progression prediction cohort**

**Baseline characteristics**
The baseline clinical and biochemical characteristics of the DKD progression cohort stratified by the progression of DKD are listed in table 1. Of the cohort, the average BMI was 26.30±4.14 kg/m$^2$ and 51.0% were men. The proportion of smokers and drinkers were 30.6% and 25.3%, respectively. In addition, the progression group had lower education and married level, and longer duration of diabetes than the patients in non-progression group. And patients in progression group had higher SBP, HbA1c and UACR than the patients in non-progression group, and higher levels of DBP, total cholesterol than the patients in non-progression group. On DKD risk score analyses, the progression group had higher baseline risk score compared with non-progression group whatever patients with low DKD risk, moderate DKD risk or high DKD risk (table 2).

**External validation results for progression prediction cohort**
Of the 4455 patients, 2504 (56.2%), 1397 (31.4%) and 554 (12.4%) were in low, moderate and high DKD risk at baseline. At the end of observation, 589 (32.6%), 531 (36.1%) and 414 (43.5%) patients in the low, moderate and high DKD risk had progressed, respectively (figure 4).
### Table 1  Baseline clinical and biochemical characteristics of the study participants stratified by the occurrence and progression of DKD

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Occurrence of DKD</th>
<th></th>
<th></th>
<th>Progression of DKD</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>No</td>
<td>Yes</td>
<td>P value</td>
<td>Total</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>2504</td>
<td>1687</td>
<td>817</td>
<td>0.059</td>
<td>4455</td>
<td>2892</td>
<td>1563</td>
</tr>
<tr>
<td>Age (years)</td>
<td>55.44±7.49</td>
<td>55.22±7.66</td>
<td>55.92±7.10</td>
<td>0.059</td>
<td>57.88±8.80</td>
<td>56.56±8.61</td>
<td>60.33±8.57</td>
</tr>
<tr>
<td>Male gender (n (%))</td>
<td>991 (54.8)</td>
<td>683 (55.2)</td>
<td>308 (53.8)</td>
<td>0.587</td>
<td>2274 (51.0)</td>
<td>1504 (52.0)</td>
<td>770 (49.3)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.10±4.00</td>
<td>25.93±3.86</td>
<td>26.44±3.86</td>
<td>0.029</td>
<td>26.30±4.14</td>
<td>26.23±4.07</td>
<td>26.43±4.27</td>
</tr>
<tr>
<td>Diabetes duration (years)</td>
<td>4 (1,8)</td>
<td>3 (1,7)</td>
<td>5 (2,9)</td>
<td>&lt;0.001</td>
<td>5 (1,10)</td>
<td>4 (1,8)</td>
<td>6 (3,11)</td>
</tr>
<tr>
<td>Education(n (%))</td>
<td>0.0021</td>
<td></td>
<td></td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illiterate</td>
<td>182 (7.3)</td>
<td>91 (5.4)</td>
<td>91 (11.1)</td>
<td>0.587</td>
<td>401 (9.0)</td>
<td>205 (7.1)</td>
<td>196 (12.5)</td>
</tr>
<tr>
<td>Literate</td>
<td>75 (3.0)</td>
<td>33 (2.0)</td>
<td>42 (5.1)</td>
<td>0.587</td>
<td>132 (3.0)</td>
<td>55 (1.9)</td>
<td>77 (4.9)</td>
</tr>
<tr>
<td>Elementary school</td>
<td>678 (27.1)</td>
<td>406 (24.1)</td>
<td>272 (33.3)</td>
<td>0.587</td>
<td>1253 (28.1)</td>
<td>727 (25.1)</td>
<td>526 (33.7)</td>
</tr>
<tr>
<td>Junior high school</td>
<td>375 (15.0)</td>
<td>260 (15.4)</td>
<td>115 (14.1)</td>
<td>0.587</td>
<td>693 (15.6)</td>
<td>462 (16.0)</td>
<td>231 (14.8)</td>
</tr>
<tr>
<td>High school</td>
<td>738 (29.5)</td>
<td>539 (32.0)</td>
<td>199 (24.4)</td>
<td>0.587</td>
<td>1228 (27.6)</td>
<td>881 (30.5)</td>
<td>347 (22.2)</td>
</tr>
<tr>
<td>College</td>
<td>400 (16.0)</td>
<td>313 (18.6)</td>
<td>87 (10.6)</td>
<td>0.587</td>
<td>649 (14.6)</td>
<td>487 (16.8)</td>
<td>162 (10.4)</td>
</tr>
<tr>
<td>No data</td>
<td>56 (2.2)</td>
<td>45 (2.7)</td>
<td>11 (1.3)</td>
<td>0.587</td>
<td>99 (2.2)</td>
<td>75 (2.6)</td>
<td>24 (1.5)</td>
</tr>
<tr>
<td>Marital status (n (%))</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Single</td>
<td>95 (3.8)</td>
<td>66 (3.9)</td>
<td>29 (3.5)</td>
<td>0.587</td>
<td>181 (4.1)</td>
<td>125 (4.3)</td>
<td>56 (3.6)</td>
</tr>
<tr>
<td>Married</td>
<td>2109 (84.2)</td>
<td>1457 (86.4)</td>
<td>652 (79.8)</td>
<td>0.587</td>
<td>3684 (82.7)</td>
<td>2422 (83.7)</td>
<td>1262 (80.7)</td>
</tr>
<tr>
<td>Divorced</td>
<td>93 (3.7)</td>
<td>57 (3.4)</td>
<td>36 (4.4)</td>
<td>0.587</td>
<td>160 (3.6)</td>
<td>105 (3.6)</td>
<td>55 (3.5)</td>
</tr>
<tr>
<td>Widow or widower</td>
<td>171 (6.8)</td>
<td>75 (4.4)</td>
<td>96 (11.8)</td>
<td>0.587</td>
<td>355 (8.0)</td>
<td>181 (6.3)</td>
<td>174 (11.1)</td>
</tr>
<tr>
<td>No data</td>
<td>36 (1.4)</td>
<td>32 (1.9)</td>
<td>4 (0.5)</td>
<td>0.587</td>
<td>75 (1.7)</td>
<td>59 (2.0)</td>
<td>16 (1.0)</td>
</tr>
<tr>
<td>Smoking (n (%))</td>
<td>769 (30.7)</td>
<td>523 (31.0)</td>
<td>246 (30.1)</td>
<td>0.710</td>
<td>1362 (30.6)</td>
<td>891 (30.8)</td>
<td>471 (30.1)</td>
</tr>
<tr>
<td>Drinking (n (%))</td>
<td>637 (25.4)</td>
<td>433 (25.7)</td>
<td>204 (25.0)</td>
<td>0.997</td>
<td>1124 (25.2)</td>
<td>743 (25.1)</td>
<td>381 (24.4)</td>
</tr>
<tr>
<td>Diabetic self-management behaviour</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet</td>
<td>5 (4.7)</td>
<td>5 (4.7)</td>
<td>5 (4.7)</td>
<td>0.576</td>
<td>5 (4.7)</td>
<td>5 (4.7)</td>
<td>5 (3.7)</td>
</tr>
<tr>
<td>Exercise</td>
<td>7 (0.7)</td>
<td>7 (1.7)</td>
<td>7 (0.7)</td>
<td>0.576</td>
<td>7 (1.7)</td>
<td>7 (0.7)</td>
<td>6.5 (0.7)</td>
</tr>
<tr>
<td>Medication</td>
<td>7 (7.7)</td>
<td>7 (7.7)</td>
<td>7 (7.7)</td>
<td>0.576</td>
<td>7 (7.7)</td>
<td>7 (7.7)</td>
<td>7 (7.7)</td>
</tr>
<tr>
<td>Monitoring</td>
<td>0 (0.7)</td>
<td>0 (0.7)</td>
<td>2 (0.7)</td>
<td>0.576</td>
<td>0 (0.7)</td>
<td>0 (0.7)</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>SBP (mm Hg)</td>
<td>129 (119,141)</td>
<td>128 (117,139)</td>
<td>133 (122,146)</td>
<td>&lt;0.001</td>
<td>133 (121,146)</td>
<td>131 (120,144)</td>
<td>136 (124,150)</td>
</tr>
<tr>
<td>DBP (mm Hg)</td>
<td>77 (70,84)</td>
<td>76 (69,83)</td>
<td>78 (71,86)</td>
<td>&lt;0.001</td>
<td>78 (71,86)</td>
<td>77 (70,85)</td>
<td>78 (71,87)</td>
</tr>
<tr>
<td>HbA1c (%)</td>
<td>7.90 (6.85,9.75)</td>
<td>7.80 (6.80,9.50)</td>
<td>8.40 (7.10,10.10)</td>
<td>&lt;0.001</td>
<td>8.2 (7.0,10.0)</td>
<td>8.0 (6.9,9.7)</td>
<td>8.5 (7.1,10.5)</td>
</tr>
<tr>
<td>TG (mg/dL)</td>
<td>113 (83,164)</td>
<td>111 (84,160)</td>
<td>120 (82,168)</td>
<td>0.094</td>
<td>123 (88,178)</td>
<td>119 (86,172)</td>
<td>133 (93,190)</td>
</tr>
</tbody>
</table>
Discrimination of the model

ROC curve showed moderate discriminative ability of predicting the progression of DKD in progression prediction cohort, and the area under the ROC was 0.637 (95% CI 0.620 to 0.654). Eighteen was selected as the optimal cut-off risk score value with a sensitivity of 65.0% (95% CI 62.6% to 67.3%) and a specificity of 57.0% (95% CI 55.2% to 58.8%).

Calibration of the model

Calibration of the model for predicting the progression of DKD was no good in our external validation cohort (Hosmer-Lemeshow \( \chi^2 = 23.663, p = 0.003 \)), and the overestimation of risk occurred in 1–10. The calibration plot in figure 5 shows that the comparison between observed and predicted the progression of DKD.

Precision of the model in predicting DKD progression

The prediction accuracy refers to the ratio of the actual number of patients whose DKD stage progressed to the number of patients predicted to progress. According to

---

**Table 1**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Occurrence of DKD</th>
<th>Total No</th>
<th>Yes</th>
<th>P value</th>
<th>Total No</th>
<th>Yes</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDL-C (mg/dL)</td>
<td>52.00±14.51</td>
<td>50.73±14.38</td>
<td>0.038</td>
<td>0.001</td>
<td>51.31±13.41</td>
<td>0.18</td>
<td>0.001</td>
</tr>
<tr>
<td>LDL-C (mg/dL)</td>
<td>99.41±28.56</td>
<td>100.10±28.60</td>
<td>0.401</td>
<td>0.001</td>
<td>101.05±29.96</td>
<td>0.18</td>
<td>0.001</td>
</tr>
<tr>
<td>eGFR (ml/min/1.73m²)</td>
<td>85.50±73.08±96.13</td>
<td>87.18±(74.00,96.90,82.76)</td>
<td>0.002</td>
<td>0.001</td>
<td>83.65±(68.88,69.36,82.01)</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>UACR (mg/g)</td>
<td>8.64±(4.50,11.98)</td>
<td>7.08±(3.58,12.17)</td>
<td>0.001</td>
<td>0.001</td>
<td>14.84±(6.57,4.17,7.22)</td>
<td>0.001</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Data are expressed as mean±SD, number (%) or median (IQR). BMI, body mass index; DBP, diastolic blood pressure; DKD, diabetic kidney disease; eGFR, estimated glomerular filtration rate; HbA1c, haemoglobin A1c; HDL-C, high-density lipoprotein-cholesterol; LDL-C, low-density lipoprotein-cholesterol; SBP, systolic blood pressure; TG, triglycerides; UACR, urinary albumin-to-creatinine ratio.

---

**Figure 2** In the DKD prediction cohort, the final DKD, non-DKD patients’ ratio stratified by DKD risk categories. DKD, diabetic kidney disease.

**Figure 3** Observed and predicted the occurrence of DKD. Grouping method: First, calculate the predicted probability of DKD for each patient according to the DKD risk scoring model, then arrange them from small to large according to the predicted probability and divide them into 10 groups based on deciles for the predicted probabilities. DKD, diabetic kidney disease.
our validation cohort, with a risk cut-off value of 18 points, the precision would be 45.0% (1016/2260). According to the model developed study,15 with a risk cut-off value of 16 points, the precision would be 42.2% (1173/2781).

DISCUSSION

External validation is a mandatory step in applying a prediction model to meaningful clinical care; the process addresses the transportability of the model.18 Taiwan has established a sound universal health insurance policy supported by the government, providing a lot of reliable data for model validation. However, not all governments have established such a comprehensive universal healthcare policy. Hence, the limited availability of patient data, the model was not validated. In this study, we evaluated the usefulness of RPM-DKD for predicting the DKD incidence and progression of DKD in patients with T2DM by assessing its discrimination, calibration and precision. The performance of the RPM-DKD predictive potential in our validation cohort was not ideal, even when the results from the validation evaluated by the model developers were promising.

The RPM-DKD has several advantages, including easy point-of-care application, simple calculation and reliance on very few variables. Nevertheless, the RPM-DKD demonstrated moderately discriminatory ability in our cohort. With that said, the RPM-DKD was not well calibrated for predicting the occurrence of DKD, and it overestimated the progression of disease. When using the same foundational thresholds established by the original RPM-DKD study, precision for predicting DKD occurrence and progression in our validation cohort was low, with values of 43.2% and 42.2%, respectively. Using the developers’ suggested thresholds resulted in inappropriate prediction of DKD in our cohort. Even at a threshold of 18 (at which the Youden’s index was maximal in our study), the precision of the model in predicting DKD progression was also low.

Overall transportability of the model was poor in our analysis, owing perhaps to the phenomenon of over-fitting. If internal validation such as bootstrap would have been performed after model developed, the phenomenon might have been foreseen.19 Furthermore, the poor external validation performance may also be closely related to the fact that our validation cohort was not ideal, even when the results from the validation evaluated by the model developers were promising.

The RPM-DKD uses cross-sectional baseline data to predict a patient’s risk of DKD 5–10 years later; it is therefore based on the assumption that there is no significant change in relevant indicators of the patient in subsequent years—a somewhat unrealistic expectation in a real world application. Since the occurrence and development of DKD is a relatively long process, metabolic indicators 5, 10 years even in preceding decades can impact subsequent outcomes. However, the model, established by logistic regression, volatility of

<table>
<thead>
<tr>
<th>Risk stage</th>
<th>Non-progression group</th>
<th>Progression group</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-high DKD risk</td>
<td>16.80±6.21</td>
<td>19.77±5.77</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Low DKD risk</td>
<td>14.48±5.61</td>
<td>17.83±5.74</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Moderate DKD risk</td>
<td>19.46±5.52</td>
<td>21.40±4.87</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>High DKD risk</td>
<td>21.68±5.22</td>
<td>22.93±5.15</td>
<td>0.005</td>
</tr>
</tbody>
</table>

DKD, diabetic kidney disease.
various parameters in the next few years was not considered, which results in the prediction performance was not high.

In our study, not all data specific to risk factors included in the RPM-DKD model were available (age, BMI, HbA1c); this may have contributed to poor model performance. In previous work, we found that HbA1c variability is an independent risk factor for nephropathy in patients with T2DM. In addition, Viazzi et al demonstrated that the variability of SBP and pulse pressure are also key influencing factors for the occurrence and development of DKD. Thus, these parameters should be strong outcome predictors for developing DKD; and yet, the RPM-DKD does not take them into account.

The RPM-DKD model cannot predict future risk of DKD in patients with T2DM aged <39 years despite early age of diagnosis being an established risk factor for developing DKD. Several groups reported an increasing incidence of youth-onset DKD. Given the earlier onset of DKD among patients with T2DM, we believe that having validated risk assessment models that include young adults may be of greater clinical use.

**Strengths and limitations of this study**

Our study has several strengths. First, our validation cohort was geographically different from the cohort used to derive the model; further, our team was not involved in model derivation, which enabled us to conduct a true independent external validation study. In addition, our cohort had an average follow-up time of more than 5 years, a duration that positioned us to effectively identify the occurrence of outcome variables.

Several limitations existed in this study. First, while the cohort was representative of a large cohort of over 4000 adults, it was geographically restricted to Taiwan. Second, most patients in our cohort had better diabetes self-management behaviours, which could potentially have affected the study results and also explain why the model overestimated risk of occurrence and progression. However, diabetes self-management is a key factor for promoting better health outcomes among patients with DKD; longitudinally, increased awareness for the DKD burden in diabetes patients might have contributed to additional self-management behaviours with a positive effect on DKD incidence and progression. In addition, the retrospective nature of our study presents an inherent limitation, although it is a simple, flexible and low-cost method to review patient data for purposes of the present analysis; in addition, we may not have successfully captured patient data related to care that may have occurred outside of our health system. Apart, the scoring model was not developed to predict the progression of DKD, but we still used this scoring model to predict the progression of DKD, which may create inapplicability. However, there is no similar scoring model that can predict the progression of DKD at present.

**Future research**

The occurrence of DKD in our study was high (31.6%), agreeing with previous study (20%–40%) conducted in China, but significantly higher than that in the USA (26.2%). Among those with diabetes, DKD prevalence varies widely between countries, with estimates ranging from 26.2% in USA to 83.6% in Tanzania. In general, the prevalence of DKD rate in Asia and Africa are higher than those in Europe and the USA. The insidious nature of type 2 diabetes and less accessibility of developing countries to healthcare contribute to a high proportion of undiagnosed patients. The late diagnosis may partially contribute to the high prevalence of DKD in Asia and Africa. In Asia, healthcare providers and nurses are not sufficiently educated about diabetes patients, about 50% lack understanding of diabetes complications, and self-care activities are suboptimal overall. Therefore, it is extremely significant to have an accurate prediction model that could assist clinicians in real-time evaluation of patient risk and implement primary and secondary preventive measures to delay progression of disease. However, the RPM-DKD established by logistic regression did not perform well. Furthermore, prediction models can become obsolete with change in population demography, better therapeutic options and care pathways, and improvement in data recording. In the future, it may be possible to build a DKD prediction model by deep learning methods in order to improve the prediction of DKD occurrence and progression; many studies have applied deep learning with proven success.

**CONCLUSION**

Our independent external validation study revealed that, in patients with T2DM, the RPM-DKD cannot accurately predict the risk of DKD occurrence and progression. The ability to accurately estimate DKD risk is critical in advancing patient care and preventing or delaying complication risk in patients with long-standing T2DM. Newer prediction models leveraging deep learning methods may prove useful for predicting risk of developing or progressing DKD.

**Acknowledgements** The authors sincerely thank all participants for their help with this study.

**Contributors** QL conceived and designed the study. ZS, KW, XY and J-L collected the epidemiological and clinical data. ZS, KW and XY drafted the manuscript. ZS, KW, XY and QL contributed to acquisition, analysis and interpretation of data. QL and JDM contributed to critical revision of the manuscript for important intellectual content. All authors read, revised and approved the final draft. QL is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Funding** This work was supported by the National Key R&D Program of China (No.2021YFYF0204809) and Key R&D Program of Hainan Province (No. ZDYF2021SHFZ236).

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were not involved in the design, conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not applicable.