



# External validation of prognostic models for chronic kidney disease among type 2 diabetes

Sigit Ari Saputro<sup>1,2</sup> · Anuchate Pattanateepapon<sup>1</sup> · Oraluck Pattanaprteep<sup>1</sup> · Wichai Aekplakorn<sup>3</sup> · Gareth J. McKay<sup>4</sup> · John Attia<sup>5</sup> · Ammarin Thakkinstian<sup>1</sup>

Received: 16 July 2021 / Accepted: 30 November 2021 / Published online: 8 January 2022  
© The Author(s) 2022

## Abstract

**Background** Various prognostic models have been derived to predict chronic kidney disease (CKD) development in type 2 diabetes (T2D). However, their generalisability and predictive performance in different populations remain largely unvalidated. This study aimed to externally validate several prognostic models of CKD in a T2D Thai cohort.

**Methods** A nationwide survey was linked with hospital databases to create a prospective cohort of patients with diabetes ( $n=3416$ ). We undertook a systematic review to identify prognostic models and traditional metrics (i.e., discrimination and calibration) to compare model performance for CKD prediction. We updated prognostic models by including additional clinical parameters to optimise model performance in the Thai setting.

**Results** Six relevant previously published models were identified. At baseline, C-statistics ranged from 0.585 (0.565–0.605) to 0.786 (0.765–0.806) for CKD and 0.657 (0.610–0.703) to 0.760 (0.705–0.816) for end-stage renal disease (ESRD). All original CKD models showed fair calibration with Observed/Expected (O/E) ratios ranging from 0.999 (0.975–1.024) to 1.009 (0.929–1.090). *Hosmer–Lemeshow* tests indicated a good fit for all models. The addition of routine clinical factors (i.e., glucose level and oral diabetes medications) enhanced model prediction by improved C-statistics of Low's of 0.114 for CKD and Elley's of 0.025 for ESRD.

**Conclusions** All models showed moderate discrimination and fair calibration. Updating models to include routine clinical factors substantially enhanced their accuracy. Low's (developed in Singapore) and Elley's model (developed in New Zealand), outperformed the other models evaluated. These models can assist clinicians to improve the risk-stratification of diabetic patients for CKD and/or ESRD in the regions settings are similar to Thailand.

---

✉ Anuchate Pattanateepapon  
anuchate.gab@mahidol.ac.th

✉ Wichai Aekplakorn  
wichai.aek@mahidol.ac.th

<sup>1</sup> Department of Clinical Epidemiology and Biostatistics, Faculty of Medicine Ramathibodi Hospital, Mahidol University, 270 Rama VI Road, Phayathai, Bangkok 10400, Thailand

<sup>2</sup> Department of Epidemiology Biostatistics Population and Health Promotion, Faculty of Public Health, Airlangga University, Surabaya 60115, Indonesia

<sup>3</sup> Department of Community Medicine, Faculty of Medicine Ramathibodi Hospital, Mahidol University, 270 Rama VI Road, Phayathai, Bangkok 10400, Thailand

<sup>4</sup> Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, UK

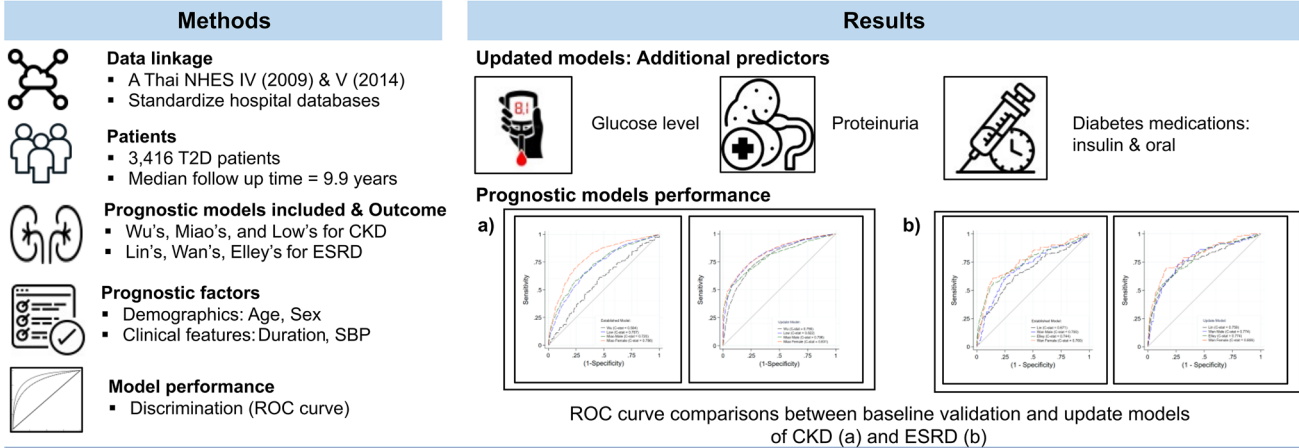
<sup>5</sup> School of Medicine and Public Health, and Hunter Medical Research Institute, University of Newcastle, New Lambton, NSW, Australia

## Graphical abstract

## External validation of prognostic models for chronic kidney diseases among type 2 diabetes



### Prospective Cohort Study



**Conclusions:** Low's (developed in Singapore) and Elley's model (developed in New Zealand), outperformed the other models.

DOI: <https://doi.org/10.1007/s40620-021-01220-w>

**Keywords** External validation · Prognostic model · Chronic kidney disease · Type 2 Diabetes

## Introduction

Chronic Kidney Disease (CKD) is a major worldwide health burden and the most common microvascular complication of type 2 diabetes (T2D) [1, 2]. In 2017, more than 840 million individuals developed CKD [3], increasing health care demand, particularly in low to middle-income countries (LMICs) [1]. In the UK and the United States, the prevalence of CKD in T2D was reported to range between 25 and 36%, of which 19% was estimated to be advanced (stages 3–5) [4]. The age-standardised global mortality of CKD due to diabetes has been estimated at 7.6 per 100,000 [5].

Early detection and treatment are beneficial in the prevention or delay of CKD progression. Despite improved screening, many CKD patients face delayed diagnosis until an advanced stage due to a lack of overt symptoms. Prognostic models for complications associated with T2D progression that incorporate clinical information systems would facilitate improved treatment allocations, healthcare management, and improve understanding of clinical research strategies [6, 7].

Currently, several prognostic equations [8–15] are available for the prediction of CKD in T2D patients, but their

generalisability remains uncertain due to limited external and independent validation, particularly in Asian populations [16, 17]. Indeed, external validation is essential and has become mandatory before implementation in clinical practice [16, 18, 19].

Despite many potential advantages, prognostic models have several shortcomings and frequently reported deficiencies [20]. Multiple models have been developed in different ethnicities [8–15, 21–29] but no single model has consistently outperformed all others in Asian populations. For instance, a study based in China performed a limited temporal internal model validation over time on the same data [10]. Most importantly, adaptation of a suitable prognostic model by ethnicity is particularly in an Asian context given that half of the ten countries affected by diabetes worldwide are Asian [4]. Furthermore, recent recommendations have proposed re-evaluation to including race/ethnicity in CKD prediction models [30].

Therefore, this study conducted external validation and improvement of previously published prognostic models of CKD and end stage renal disease (ESRD) in Thai T2D patients.

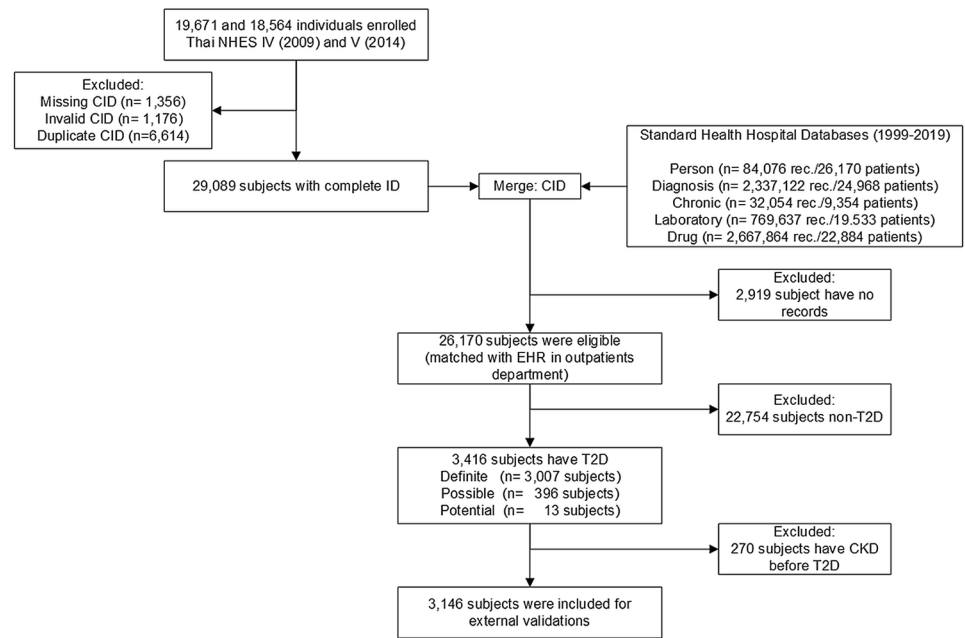
**Table 1** Characteristics of prognostic studies that were used for external validations

Characteristics	Thai NHES				Prognostic score			
	Wu et al. [10]	Miao et al. [9]	Low et al. [8]	Lin et al. [12]	Wan et al. [13]	Elley et al. [11]		
Study design	Cross-sectional with updated retrospective data	Retrospective cohort	Retrospective cohort	Retrospective cohort	Retrospective cohort	Retrospective cohort		
Country	Thailand	China	Singapore	Taiwan	Hong Kong	New Zealand		
Outcome	CKD-ESRD	CKD	CKD	ESRD	ESRD	ESRD		
Sample size	3416	11,771	1582	24,104	116,509	25,736		
Incidence, n (%)	CKD:1,383 (43.9) ESRD: 186 (5.9)	77 (0.7)	679 (42.9)	1215 (5.04)	238 (0.4)	637 (2.5)		
Study settings	Community – hospital-based	Community – hospital-based	Hospital-based	Hospital-based	Hospital-based	Hospital-based		
Study period included	1999 – 2019	2014	2002 – 2014	2001 – 2011	2010 – 2015	1988–2010		
Model type	Logistic	Cox	Logistic	Cox	Cox	Cox		
Handling missing data	Excluded	Multiple imputation	Excluded	Multiple imputation	MICE	Excluded		
Internal validation	n/a	Split sample	Split sample	Split sample Bootstrapping	Split sample	Yes (temporal external validation)		
External validation	Yes (standalone external validation)	No	No	No	No	Yes (temporal external validation)		
Predicted time, years	n/a	5, 10, 20	–	3, 5, 8	5	5		
Type 2 diabetes ascertainment	Physical examination at survey (FPG ≥ 126 mg/dL), linked to standard hospital databases (ICD-10) and Laboratory follow up	Medical records, FPG ≥ 7.0 mmol/L, self-report diabetes, or ICD-IX (code 250.X)	Medical records, and standardized questionnaire	Medical records, classification based ADA, 2006 (i.e., FPG ≥ 7.0 or OGTT ≥ 11.1 mmol/L)	Medical records based on ICD-9 version II (code T90)	Medical records, classification based WHO, 1998 (i.e., FPG ≥ 7.0 or OGTT ≥ 11.1 mmol/L)		
Outcome ascertainment	ICD-10 code diagnosis, laboratory follow up, death certificate, and physical examination at survey	Persistent albuminuria, medical records review	eGFR- MDRD: < 60 mL/min/1.73 m <sup>2</sup>	ICD-9 (585.x), doubling S-Cr > 2.26 mg/dL, RRT, renal death, or eGFR- MDRD < 15 mL/min/1.73 m <sup>2</sup>	ICD-9 (250.3x, 585.x, 586.x), eGFR < 15 mL/min/1.73 m <sup>2</sup>	Dialysis, RRT, or renal death based on ICD-IX and ICD-10 codes		
Number of prognostic factors	n/a	8 (Male); 6 (Female)	6	11	10 (Male); 11 (Female)	10 (Equation IV)		

Table 1 (continued)

Characteristics	Thai NHES	Prognostic score	Wu et al. [10]	Miao et al. [9]	Low et al. [8]	Lin et al. [12]	Wan et al. [13]	Elley et al. [11]
Prognostic factors in the final equation	n/a	Sex, BMI, SBP, diabetes duration	Male: age, BMI, Cr, HDL, location, HT/DLP, DR, diet control/physical activity Female: age, Cr, HDL, location, HT/DLP, DR	Male: age, BMI, Cr, HDL, location, HT/DLP, DR, diet control/physical activity Female: age, Cr, HDL, location, HT/DLP, DR	Age, HbA1c, SBP, UACR, eGFR, LDL-C	Age at initial, Sex, Age on set, Cr, HbA1c, SBP, DR, albuminuria, DM drug, anti-HT drug * SBP/DBP, HLD drug * TC	Male: age, smoking, DR, anti-HT drug, oral DM drug, insulin, HbA1c, SBP, DBP, UACR, eGFR Female: age, duration, anti-HT drug, oral DM drug, insulin, BMI, HbA1c, SBP, DBP, UACR, eGFR	Sex, ethnicity, age on set, duration, Cr, albuminuria, SBP, HbA1c, smoking, CVD history
C-statistic in development phase	n/a	0.713 (0.692 – 0.734)	Male: 0.840 (0.800 – 0.880) Female: 0.800 (0.740 – 0.860)	Male: 0.840 (0.800 – 0.880) Female: 0.800 (0.740 – 0.860)	0.830 (0.790 – 0.870)	0.920 (0.900 – 0.930)	Male: 0.866 (0.849 – 0.882) Female: 0.862 (0.845 – 0.880)	0.914 (0.881 – 0.937)
Calibration	n/a	–	Fair calibration plot	Fair calibration plot	HL- $\chi^2$ ( $p$ value = 0.986)	Fair calibration plot	–	HL- $\chi^2$ ( $p$ value = 0.663)

ADA American Diabetes Association, BMI body mass index, CKD chronic kidney disease, Cr creatinine, CVD cardiovascular disease, DBP diastolic blood pressure, SBP systolic blood pressure, DLP dyslipidaemia, DM diabetic mellitus, DR diabetic retinopathy, eGFR estimated glomerular filtration rate, ESKD end stage renal disease, FPG fasting plasma glucose, HbA1c glycated haemoglobin, HDL-C high density lipoprotein cholesterol, HL Hosmer-Lemeshow chi square test, HLD hyperlipidaemia, HT hypertension, ICD International classifications of diseases, ICPC International classification of primary care, LDL-C low density lipoprotein cholesterol, MDRD Modification of diet in renal disease, MICE multiple imputation chained equations, n/a not appropriate, NHES National health examination survey, OGTT oral glucose tolerance test, RRT renal replacement therapy, TC total cholesterol, UACR urine albumin creatinine ratio, WHO World Health Organization

**Fig. 1** Flowchart for participant inclusion

## Methods

We adhered to the TRIPOD guidelines for the development and validation of a clinical prediction score [31, 32]. We focused on external validation of existing models of CKD-ESRD risk predictions in T2D, supplemented with the addition of routine clinical factors to potentially increase the discriminatory power in our local population [18].

We first identified previous prognostic models by performing a systematic review and meta-analysis (SR/MAs), see Figure S1. We selected prognostic models if they: (1) had been internally or externally validated; (2) reported moderate to excellent discrimination of C-statistics, i.e.,  $\geq 0.70$ . We identified six studies that met the inclusion criteria for CKD [8–10] and ESRD [11–13] (Table 1).

## Study design and data sources

Data from the Thailand National Health Examination Survey (Thai-NHES) and the standard health databases version 2.4 2019 edition (<http://spd.moph.go.th/healthdata/>) were used for model validation. The NHES IV and V were population-based cross-sectional surveys conducted in 2009 and 2014, respectively. These surveys captured: health interviews, physical examination, nutrition assessment, and health-related behaviours [33]. Briefly, a multi-stage sampling of adult subjects from the regions, provinces, and districts across the country was used [34, 35].

The standard health databases included medical service records from hospitals, mostly under the direction of the Ministry of Public Health. They comprised a set of tables of all transactions from outpatient and inpatient services

for each individual; of 43 files available, only the six that were related to outpatient services (i.e., Person, Diagnosis, Chronic, Drug, Laboratory, and Death) were used for this study.

## Settings and participants

A total of 19,671 and 18,564 participants were de-identified from NHES IV-V, respectively; removal of duplicates and missing or invalid citizen identification (CID) resulted in 29,089 participants remaining, see Fig. 1. These were linked with the standard hospital health databases (1999–2019) using an encrypted CID to construct the initial sampling frame, leaving a total of 26,170 participants.

We confirmed T2D status based on self-report, medication use, and/or pathology tests (Fasting Plasma Glucose (FPG)  $\geq 126$  mg/dL or HbA1c  $\geq 6.5\%$ ). We excluded type 1 diabetes (T1D) with age at onset less than 30 years with severe insulin treatment. There were 3416 participants with identified T2D, of whom 270 (7.9%) were excluded on the basis that CKD was diagnosed prior to T2D, leaving a total of 3146 participants. Of these, 3014 (10.4%) participated in both NHES IV-V, with 402 newly diagnosed participants identified after the survey, see Fig. 1. These participants were followed up from 1999 to October 31st, 2019.

## Outcomes

The primary study outcomes included diabetic nephropathy (CKD stage 3–5) based on the International Classification of Disease, Tenth Edition (ICD-X), which was confirmed by estimated glomerular filtration rate (eGFR)  $< 60$  mL/

min/1.73m<sup>2</sup> measured within 3 months before and after diagnosis, see Table S1. ESRD (CKD stage 5) was defined as eGFR < 15 mL/min/1.73m<sup>2</sup>, or dialysis identified by ICD-X code diagnosis. eGFR was based on the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) formula [36].

### Established prognostic factors

We focused on prognostic factors identified through our systematic review, including demographics (age, sex, education, income, and area of residence), biomarkers, comorbidities, medication usage, and clinical features; the latter included diabetes duration, body mass index (BMI; kg/m<sup>2</sup>), waist and hip circumference (cm), systolic/diastolic (SBP/DBP) blood pressure (mmHg), pulse (beat/min), smoking, alcohol consumption, dietary control measures, physical activity, dyslipidaemia, hypertension, and family history of diabetes (FHD, presence of T2D in 1<sup>st</sup>-degree relatives). Biomarkers included lipid profile (i.e., high-density lipoprotein (HDL), low-density lipoprotein (LDL), triglycerides (TG), total cholesterol (TC) in mg/dL, FPG (mg/dL), haemoglobin (g/dL)) and dipstick proteinuria. Comorbidities included a history of cardiovascular disease (CVD) and stroke. CVD was defined by self-report, clinical diagnosis or receipt of treatment for coronary heart disease. Medications recorded included oral-diabetic, blood pressure or cholesterol-lowering drugs.

We included clinical data associated with diabetic complications (i.e., retinopathy, stroke, and composite CVD's) based on ICD-X diagnostic codes (Table S1), laboratory follow-up, medication treatment (Table S2), or death certification (based on ICD-X).

Hypertension was defined as SBP ≥ 140 or DBP ≥ 90 mmHg or use of anti-hypertensive medication. Dyslipidaemia was defined as HDL ≤ 40 mg/dL, or LDL, TG and TC levels ≥ 130, ≥ 130, and ≥ 200, mg/dL respectively, according to ATP-III guidelines [37].

All factors were included according to their definitions in the original studies (Table S3–S4).

### Statistical analysis

Descriptive statistics for predictor variables were summarised as mean (± standard deviation) or median (interquartile range) for continuous variables or frequency (percentage) for categorical variables. Participant characteristics were compared between groups using Chi-Square or Fisher's Exact test, where appropriate for categorical variables, and one-way ANOVA or Kruskal Wallis for continuous variables. The predictors which were missing ranged from only 0.1% (n = 3) to 5.8% (n = 199). Therefore, a complete case analysis was applied for the whole analyses.

We evaluated prognostic models originally derived by logistic [8, 10] or Cox regression models [9, 11–13] that

were identified from our systematic review (PROSPERO: CRD42018105287). Prognostic scores were calculated according to the published regression formulae using the coefficient and intercept or baseline hazard, see Table S4.

External validation was undertaken in accordance with guidelines for the validation and interpretation of risk prediction models [18, 19]. In brief, we evaluated model performance through comparisons between the original published equation and models that included additional adjustment (e.g., intercept, regression coefficients) for other potential predictors, see Appendix [18, 38–40].

Briefly, model performance was evaluated as follows [7]. Discrimination was assessed by concordance of C-statistics, area under receiver operator characteristic curves (AUROC) [41], and 95% confidence intervals (CI's). Calibration, i.e., the closeness between the observed and predicted values, was assessed using the *Hosmer–Lemeshow* goodness-of-fit test, the observed to expected (O/E) ratios with 95% CI, and calibration plots. We also used global heuristic shrinkage factors and penalised regression to address issues of over-optimism in updated prognostic models [39, 42, 43].

All statistical analyses were conducted using STATA version 16.0. A two-sided p-value less than 0.05 was considered significant.

## Results

### Characteristics of prognostic models

We identified a total of 6 prognostic studies for CKD-ESRD in T2D patients; see PRISMA flow diagram in Figure S1. Of these, two [8, 10] and four [9, 11–13] applied logistic and Cox regressions, respectively (see Table S4).

Five [8–10, 12, 13] models were developed in Asia and one in New Zealand [11]. Only two [10, 11] models had been externally validated in either a Chinese or New Zealand population. Five [8–13] studies used hospital-based data. The mean age of T2D subjects ranged from 55.4 to 62.9 years with study size ranging between 1582 and 116,509. Five [8–10, 12, 13] studies performed internal validation by splitting samples for discovery and validation, and three [9, 12, 13] applied multiple imputation to account for missing data (see Table 1).

T2D was characterized on the basis of a FPG ≥ 7.0 mmol/L in four [9–12] studies, or medical record review in the remaining two studies [8, 13]. Identification of CKD was mainly based on eGFR and ICD-X codes. The number of prognostic factors included in each model varied between 4 and 11 and included age, sex, SBP, creatinine, and diabetes duration as common predictor variables.

**Table 2** Baseline characteristics of T2D in Thailand NHES IV-V

Variables	Missing <i>n</i> (%)	All patients ( <i>n</i> = 3,416)	CKD groups				<i>P</i>
			Normal ( <i>n</i> = 2884)	Stage 3 ( <i>n</i> = 470)	Stage 4 ( <i>n</i> = 53)	Stage 5 ( <i>n</i> = 9)	
Demographic and socio-economic status							
Age, years	–	56.6 (12.4)	54.6 (12.0)	67.8 (8.6)	66.4 (8.0)	64.3 (7.7)	<0.001 <sup>‡</sup>
Age at diabetes onset, years	–	60.0 (12.3)	58.2 (11.9)	70.0 (9.9)	67.4 (9.1)	65.7 (7.4)	<0.001 <sup>‡</sup>
Sex							
Male	–	1360(39.8)	1143(39.6)	189(40.2)	21(39.6)	7(77.8)	0.1 <sup>¶</sup>
Female	–	2056(60.2)	1741(60.4)	281(59.8)	32(60.4)	2(22.2)	
Education							
No formal	–	226(6.6)	168(5.8)	55(11.7)	3(5.7)	0(0.0)	<0.001 <sup>§</sup>
Primary	–	2415(70.7)	1999(69.3)	364(77.4)	44(83.0)	8(88.9)	
Secondary	–	543(15.9)	508(17.6)	32(6.8)	3(5.7)	0(0.0)	
University	–	232(6.8)	209(7.2)	19(4.0)	3(5.7)	1(11.1)	
Personal income/month (THB)							
< 5000	–	697(20.4)	644(22.3)	44(9.4)	7(13.2)	2(22.2)	<0.001 <sup>§</sup>
5000 – 10,000	–	545(16.0)	525(18.2)	19(4.0)	1(1.9)	0(0.0)	
10,000 – 25,000	–	450(13.2)	441(15.3)	8(1.7)	1(1.9)	0(0.0)	
≥ 25,000	–	145(4.2)	139(4.8)	4(0.9)	2(3.8)	0(0.0)	
Not answered	–	1579(46.2)	1135(39.4)	395(84.0)	42(79.2)	7(77.8)	
Area of residence							
Rural	–	1608(47.1)	1365(47.3)	213(45.3)	25(47.2)	5(55.6)	0.8 <sup>§</sup>
Urban	–	1808(52.9)	1519(52.7)	257(54.7)	28(52.8)	4(44.4)	
Clinical features							
Diabetic duration, years		5.7 (2.6–10.1)	5.5 (2.6–10.1)	7.0 (3.2–13.2)	9.0 (4.8–17.3)	10.4 (4.9–15.3)	<0.001 <sup>‡</sup>
BMI, kg/m <sup>2</sup>	21 (0.6)	26.4 (4.7)	26.6 (4.7)	25.5 (4.3)	25.3 (4.6)	24.3 (7.9)	<0.001 <sup>‡</sup>
Waist circumference, cm	9 (0.2)	87.7 (11.2)	87.8 (11.3)	87.1 (10.7)	87.1 (12.7)	82.4 (14.0)	0.3 <sup>†</sup>
Hip circumference, cm	15 (0.4)	97.2 (9.5)	97.5 (9.5)	95.5 (9.2)	95.5 (11.2)	93.4 (15.1)	<0.001 <sup>‡</sup>
Blood pressure							
SBP, mmHg	3 (0.09)	132.6 (19.8)	131.9 (19.3)	136.0 (20.8)	141.1 (29.1)	142.0 (28.6)	<0.001 <sup>‡</sup>
DBP, mmHg	3 (0.09)	78.9 (11.3)	79.4 (11.2)	76.0 (10.8)	79.0 (15.6)	74.8 (10.7)	<0.001 <sup>‡</sup>
Pulse, beat/min	4 (0.12)	78.5 (13.1)	78.5 (12.7)	78.5 (15.1)	77.2 (16.3)	79.2 (8.3)	0.6 <sup>‡</sup>
Smoking status							
Non-smoker	–	2375(69.5)	2014(69.8)	324(68.9)	33(62.3)	4(44.4)	<0.001 <sup>¶</sup>
Current smoker	–	495(14.5)	440(15.3)	45(9.6)	8(15.1)	2(22.2)	
Past smoker	–	546(16.0)	430(14.9)	101(21.5)	12(22.6)	3(33.3)	
Alcohol drinking							
No	–	2367(69.3)	1927(66.8)	389(82.8)	44(83.0)	7(77.8)	<0.001 <sup>¶</sup>
Yes	–	1049(30.7)	957(33.2)	81(17.2)	9(17.0)	2(22.2)	
Dietary control							
No	–	2228(65.2)	1885 (65.4)	315 (68.9)	33 (62.3)	6 (66.7)	0.003 <sup>¶</sup>
Yes	–	1188(34.8)	999 (34.6)	45 (9.6)	8 (15.1)	3 (33.3)	

**Table 2** (continued)

Variables	Missing <i>n</i> (%)	All patients ( <i>n</i> = 3,416)	CKD groups				<i>P</i>
			Normal ( <i>n</i> = 2884)	Stage 3 ( <i>n</i> = 470)	Stage 4 ( <i>n</i> = 53)	Stage 5 ( <i>n</i> = 9)	
Physical activity	25 (0.7)						
Low	–	770(22.7)	615(21.5)	128(27.5)	25(47.2)	2(25.0)	< 0.001 <sup>§</sup>
Moderate	–	1208(35.6)	1006(35.1)	179(38.5)	19(35.8)	4(50.0)	
High	–	1413(41.7)	1244(43.4)	158(34.0)	9(17.0)	2(25.0)	
FHD in 1st degree relatives							
No	–	2512(73.5)	2029(70.4)	428(91.1)	47(88.7)	8(88.9)	< 0.001 <sup>¶</sup>
Yes	–	904(26.5)	855(29.6)	42(8.9)	6(11.3)	1(11.1)	
Dyslipidaemia							
No	–	505 (14.8)	452 (15.7)	49 (10.4)	3 (5.7)	1 (11.1)	0.006 <sup>¶</sup>
Yes	–	2911 (85.2)	2432 (84.3)	421 (89.6)	50 (94.3)	8 (88.9)	
Presence of hypertension							
No	–	1630 (47.7)	1465 (50.8)	149 (31.7)	14 (26.4)	2 (22.2)	< 0.001 <sup>§</sup>
Yes	–	1786 (52.3)	1419 (49.2)	321 (68.3)	39 (73.6)	7 (77.8)	
Biomarkers							
FPG, mg/dL	199 (5.8)	138.3 (61.5)	139.1 (61.2)	133.4 (60.8)	140.6 (82.8)	132.5 (60.2)	0.4 <sup>‡</sup>
Lipid profile							
HDL-C, mg/dL	80 (2.3)	44.1 (11.2)	44.7 (11.2)	41.3 (10.6)	37.4 (9.2)	40.8 (20.6)	< 0.001 <sup>‡</sup>
LDL-C, mg/dL	94 (2.7)	134.1 (41.5)	135.5 (41.3)	126.7 (40.8)	132.9 (52.2)	94.0 (23.6)	< 0.001 <sup>‡</sup>
TG, mg/dL	95 (2.8)	158.6 (112.0–227.7)	155.1 (110.8–225.0)	170.6 (119.6–234.8)	181.6 (141.8–244.5)	124.0 (92.0–197.6)	< 0.001 <sup>‡</sup>
TC, mg/dL	94 (2.7)	213.2 (47.7)	214.4 (47.7)	206.8 (46.5)	212.0 (55.6)	160.5 (33.8)	< 0.001 <sup>†</sup>
Serum creatinine, mg/dL	102 (2.9)	0.9 (0.7–1.0)	0.8 (0.7–1.0)	1.3 (1.1–1.5)	2.2 (2.0–2.7)	5.9 (4.7–11.7)	< 0.001 <sup>‡</sup>
eGFR, mL/min/1.73 m <sup>2</sup>	102 (2.9)	84.6 (24.0)	92.2 (17.5)	48.4 (9.0)	24.5 (5.3)	8.2 (4.4)	< 0.001 <sup>‡</sup>
Blood Haemoglobin, g/dL	74 (2.1)	13.1 (1.7)	13.3 (1.6)	12.4 (1.7)	11.4 (2.1)	10.7 (1.7)	< 0.001 <sup>‡</sup>
Dipstick proteinuria	102 (2.9)						
> Trace	–	2874 (86.7)	2496 (89.0)	346 (76.5)	31 (60.8)	1 (14.3)	< 0.001 <sup>§</sup>
> 1 g/dL	–	395 (11.9)	282 (10.1)	94 (20.8)	15 (29.4)	4 (57.1)	
> 3 g/dL	–	45 (1.4)	26 (0.9)	12 (2.7)	5 (9.8)	2 (28.6)	
Drug Usage							
Oral diabetic drug							
No	–	2194(64.2)	1961(68.0)	210(44.7)	19(35.8)	4(44.4)	< 0.001 <sup>¶</sup>
Yes	–	1222(35.8)	923(32.0)	260(55.3)	34(64.2)	5(55.6)	
Insulin treatment							
No	–	3136(91.8)	2704(93.8)	394(83.8)	31(58.5)	7(77.8)	< 0.001 <sup>§</sup>
Yes	–	280(8.2)	180(6.2)	76(16.2)	22(41.5)	2(22.2)	
Blood-pressure lowering drug							
No	–	2215(64.8)	1983(68.8)	204(43.4)	25(47.2)	3(33.3)	< 0.001 <sup>¶</sup>
Yes	–	1201(35.2)	901(31.2)	266(56.6)	28(52.8)	6(66.7)	
Cholesterol-lowering drug							
No	–	2633(77.1)	2271(78.7)	323(68.7)	36(67.9)	3(33.3)	< 0.001 <sup>¶</sup>
Yes	–	783(22.9)	613(21.3)	147(31.3)	17(32.1)	6(66.7)	



**Table 2** (continued)

Variables	Missing <i>n</i> (%)	All patients ( <i>n</i> = 3,416)	CKD groups				<i>P</i>
			Normal ( <i>n</i> = 2884)	Stage 3 ( <i>n</i> = 470)	Stage 4 ( <i>n</i> = 53)	Stage 5 ( <i>n</i> = 9)	
<b>NSAIDs</b>							
No	–	1223(35.8)	1032(35.8)	167(35.5)	22(41.5)	2(22.2)	0.7 <sup>¶</sup>
Yes	–	2193(64.2)	1852(64.2)	303(64.5)	31(58.5)	7(77.8)	
<b>Comorbidities</b>							
<b>Diabetic retinopathy</b>							
No	–	2921 (85.5)	2464 (85.4)	401 (85.3)	47 (88.7)	9 (100.0)	<0.6 <sup>¶</sup>
Yes	–	495 (14.5)	420 (14.6)	69 (14.7)	6 (11.3)	0 (0.0)	
<b>History of CHD</b>							
No	–	3279 (96.0)	2792 (96.8)	430 (91.5)	48 (90.6)	9 (100.0)	<0.001 <sup>§</sup>
Yes	–	137 (4.0)	92 (3.2)	40 (8.5)	5 (9.4)	0 (0.0)	
<b>History of stroke</b>							
No	–	3323 (97.3)	2815 (97.6)	449 (95.5)	51 (96.2)	8 (88.9)	0.02 <sup>§</sup>
Yes	–	93 (2.7)	69 (2.4)	21 (4.5)	2 (3.8)	1 (11.1)	

Continuous value is presented as mean (SD) or median (IQ), while categorical was showed as numbers (%) when appropriate

*BMI* body mass index, *CHD* coronary heart diseases, *DBP* diastolic blood pressure, *e-GFR* estimated glomerular filtration rate, *FHD* family history of diabetes, *FPG* fasting plasma glucose, *HDL-C* high-density lipoprotein cholesterol, *LDL-C* low-density lipoprotein cholesterol, *NSAIDs* nonsteroidal anti-inflammatory drugs, *SBP* systolic blood pressure, *TC* total cholesterol, *TG* triglycerides *THB* Thai baht rate (฿)

Comparisons (*p* value) were obtained by

<sup>¶</sup>Pearson's Chi Square

<sup>§</sup>Fisher's Exact Chi Square Test

<sup>†</sup>One-way ANOVA otherwise

<sup>‡</sup>ANOVA Kruskal Wallis

These models had fair to good calibration, and discrimination C-statistics ranged between 0.713 [10] and 0.920 [12].

### NHES population characteristics

The T2D cohort included 3,416 participants with a median diabetes duration and follow up time of 5.7 (IQR: 2.6–10.1) and 9.9 (IQR: 6.8–12.7) years, respectively, see Table 1. Of these, 1383 and 186 participants developed CKD and ESRD with an incidence (95% CI) of 43.9% (42.2%, 45.7%) and 5.9% (5.1%, 6.8%), respectively; 704 (22.3%) and 495 (14.5%) developed CVD and retinopathy, and 420 (12.3%) died from any cause.

Baseline characteristics of T2D patients are described in Table 2. The mean (SD) age was 56.6 (12.4) years, and 60.2% were female. The mean age at diabetes onset was 60.0 (12.3) years, and 26.5% of patients had a first degree relative with diabetes. Mean BMI was 26.4 (4.7) kg/m<sup>2</sup>, and the presence of hypertension and dyslipidaemia was 52.3% and 85.2%, respectively.

A total of 1,222 (35.8%), 280 (8.2%), and 1,188 (34.8%) participants were undergoing treatment for diabetes, including oral diabetic medications, insulin, or diet-control, respectively. In general, all prognostic factors including

demographics, socioeconomic status, clinical features, biomarkers, treatments, and complications demonstrated significant differences between CKD stages 3–5 (Table 2).

### Participant characteristics comparisons

Participants in our study were slightly younger with fewer males (39.8% vs 43.7%–56.2%) compared to the other six CKD-ESRD studies (Table S5). Mean diabetes duration, BMI, serum creatinine, eGFR and SBP-DBP for our cohort fell within the range reported across the various models but the prevalence of dyslipidemia and hypertension was much higher among our participants. Our cohort had lower FPG and HDL-C, but higher lipid levels (i.e., LDL-C, TG and TC). Moreover, the percentages of anti-hypertensive, anti-hyperlipidaemic and oral diabetic medications were lower than for other reported models.

CKD incidence in our study was similar to that reported by Low and colleagues [8] (i.e., 43.9 vs 42.9%), but much higher than that reported in the remaining studies [9, 10], which ranged from 0.7 to 12.3%. The incidence of ESRD in the study by Lin et al. [12] was comparable to our study (5.04% vs 5.90%), but much higher than the other two studies that reported it [11, 13] (0.4% and 2.5%), see Table 1.

**Table 3** Details of external validation performance for CKD-ESRD models

Study	Derivative		Prognostic factors		External validation								
	C-statistics (95% CI)		n <sub>1</sub>	n <sub>2</sub>	Method		Goodness of fit		O/E ratio (95% CI)	Brier Score	S <sub>VH</sub>	Penalty (%)	
	Variables missing/added in the equation		equation		C-statistics (95% CI)	C-statistics (95% CI)	H-L	df	P				
<b>Part A: original prognostic model</b>													
<b>Chronic kidney disease (CKD)</b>													
Wu et al. [10]	0.713 (0.692–0.734)	4	–	–	M <sub>0</sub>	0.585 (0.565, 0.605)	7.59	8	0.4	0.999 (0.975, 1.024)	0.241	0.985	1.48
		4	–	–	M <sub>1</sub>	–	–	–	–	–	–	–	–
		4	–	–	M <sub>2</sub>	0.585 (0.565, 0.605)	7.59	8	0.4	0.999 (0.975, 1.024)	0.241	0.985	1.48
		4	–	–	M <sub>3</sub>	0.636 (0.616, 0.656)	6.46	8	0.5	1.002 (0.978, 1.026)	0.231	0.994	0.53
		6	FPG groups * oral diabetic drugs <sup>†</sup>	–	M <sub>4</sub>	0.790 (0.774, 0.806)	16.42	8	0.09	1.003 (0.953, 1.054)	0.185	0.987	1.21
		4	–	–	M <sub>5</sub>	0.638 (0.618, 0.657)	2.88	8	0.9	1.001 (0.985, 1.017)	0.231	0.994	4.65
		4	–	–	M <sub>6</sub>	0.632 (0.612, 0.651)	3.94	8	0.7	1.002 (0.985, 1.020)	0.232	0.994	2.70
Miao et al. [9]	0.800 (0.740–0.860)	8	–	–	M <sub>0</sub>	0.720 (0.691, 0.749)	14.23	8	0.07	1.003 (0.933, 1.074)	0.211	0.994	0.55
	Male equation	8	–	–	M <sub>1</sub>	–	–	–	–	–	–	–	–
		8	–	–	M <sub>2</sub>	0.720 (0.691, 0.749)	14.23	8	0.07	1.003 (0.933, 1.074)	0.211	0.994	0.55
		8	–	–	M <sub>3</sub>	0.794 (0.768, 0.820)	49.82	8	<0.001	1.056 (0.859, 1.253)	0.179	0.997	0.27
		10	FPG groups, oral diabetic drugs <sup>†</sup>	–	M <sub>4</sub>	0.796 (0.769, 0.822)	42.74	8	<0.001	1.052 (0.868, 1.235)	0.178	0.997	0.28
		8	–	–	M <sub>5</sub>	0.826 (0.802, 0.850)	28.32	8	<0.001	1.004 (0.849, 1.159)	0.166	0.982	1.79
		8	–	–	M <sub>6</sub>	0.824 (0.800, 0.849)	40.80	8	<0.001	1.125 (0.860, 1.389)	0.165	0.988	1.13
	0.840 (0.800–0.880)	6	–	–	M <sub>0</sub>	0.786 (0.765, 0.806)	21.55	8	0.006	1.009 (0.929, 1.090)	0.187	0.997	0.21
	Female equation	6	–	–	M <sub>1</sub>	–	–	–	–	–	–	–	–
		6	–	–	M <sub>2</sub>	0.786 (0.765, 0.806)	21.55	8	0.006	1.009 (0.929, 1.090)	0.187	0.997	0.21
		6	–	–	M <sub>3</sub>	0.816 (0.797, 0.836)	26.05	8	<0.001	1.040 (0.919, 1.161)	0.169	0.998	0.15
		8	FPG groups * oral Diabetic drugs <sup>†</sup>	–	M <sub>4</sub>	0.831 (0.812, 0.851)	37.45	8	<0.001	1.036 (0.917, 1.156)	0.162	0.998	0.15
		6	–	–	M <sub>5</sub>	0.821 (0.802, 0.841)	34.26	8	<0.001	1.008 (0.849, 1.167)	0.167	0.990	0.93
		6	–	–	M <sub>6</sub>	0.820 (0.801, 0.840)	32.19	8	<0.001	1.004 (0.849, 1.159)	0.167	0.995	0.46
Low et al. [8]	0.830 (0.790–0.870)	6	4	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>0</sub>	0.707 (0.689, 0.726)	14.72	8	0.06	1.000 (0.956, 1.044)	0.216	0.997	0.25
		4	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	–	M <sub>1</sub>	0.707 (0.689, 0.726)	14.72	8	0.06	1.000 (0.956, 1.044)	0.216	0.997	0.25
		4	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	–	M <sub>2</sub>	0.708 (0.689, 0.730)	14.72	8	0.06	1.000 (0.956, 1.044)	0.216	0.997	0.25
		4	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	–	M <sub>3</sub>	0.803 (0.787, 0.819)	69.16	8	<0.001	1.013 (0.858, 1.169)	0.177	0.999	0.10
		6	FPG groups * Oral Diabetic drugs <sup>†</sup>	–	M <sub>4</sub>	0.822 (0.806, 0.837)	45.13	8	<0.001	1.016 (0.875, 1.157)	0.168	0.994	0.53
		4	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	–	M <sub>5</sub>	0.810 (0.795, 0.826)	76.69	8	<0.001	1.010 (0.836, 1.183)	0.172	0.996	0.39
		4	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	–	M <sub>6</sub>	0.810 (0.795, 0.826)	76.69	8	<0.001	1.010 (0.836, 1.183)	0.172	0.996	0.39

**Table 3** (continued)

Study	Derivative	Prognostic factors		External validation											
		C-statistics (95% CI)	n <sub>1</sub>	n <sub>2</sub>	Variables missing/added in the equation	Method		Goodness of fit	O/E ratio (95% CI)	Brier Score	S <sub>WH</sub>	Penalty (%)			
						C-statistics (95% CI)	H–L						df	P	
<b>End stage renal disease (ESRD)</b>															
Lin et al. [12]	0.920 (0.900–0.930)	11	9	HbA1c <sup>‡</sup> , SBP (SD) <sup>‡</sup>	M <sub>0</sub>	0.671 (0.626, 0.717)	17.06	8	0.02	1.026 (0.890, 1.161)	0.052	0.988	1.15		
		9	9	HbA1c <sup>‡</sup> , SBP (SD) <sup>‡</sup>	M <sub>1</sub>	–	–	–	–	–	–	–	–		
		9	9	HbA1c <sup>‡</sup> , SBP (SD) <sup>‡</sup>	M <sub>2</sub>	0.671 (0.626, 0.717)	17.06	8	0.02	1.026 (0.890, 1.161)	0.052	0.988	1.15		
		9	9	HbA1c <sup>‡</sup> , SBP (SD) <sup>‡</sup>	M <sub>3</sub>	0.747 (0.704, 0.790)	15.75	8	0.04	0.986 (0.872, 1.099)	0.051	0.993	0.66		
		11	11	FPG <sup>†</sup> , BMI <sup>†</sup>	M <sub>4</sub>	0.759 (0.716, 0.801)	7.06	8	0.5	0.973 (0.896, 1.051)	0.052	0.993	0.63		
		9	9	HbA1c <sup>‡</sup> , SBP (SD) <sup>‡</sup>	M <sub>5</sub>	0.774 (0.734, 0.812)	11.12	8	0.2	0.973 (0.863, 1.083)	0.048	0.876	12.31		
		9	9	HbA1c <sup>‡</sup> , SBP (SD) <sup>‡</sup>	M <sub>6</sub>	0.756 (0.714, 0.798)	5.78	8	0.6	0.976 (0.896, 1.056)	0.049	0.949	5.06		
		Wan et al. [13]	0.866 (0.849–0.882)	11	9	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>0</sub>	0.700 (0.639, 0.761)	10.56	8	0.2	1.014 (0.831, 1.197)	0.064	0.975	2.48
				9	9	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>1</sub>	–	–	–	–	–	–	–	–
				9	9	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>2</sub>	0.700 (0.639, 0.761)	10.56	8	0.2	1.014 (0.831, 1.197)	0.064	0.975	2.48
				9	9	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>3</sub>	0.746 (0.690, 0.803)	12.82	8	0.1	0.977 (0.818, 1.135)	0.062	0.984	1.53
Female equation	0.862 (0.845–0.880)	12	12	Albuminuria <sup>†</sup> , FPG <sup>†</sup> , LDL-C <sup>†</sup>	M <sub>4</sub>	0.774 (0.717, 0.830)	7.40	8	0.4	0.998 (0.838, 1.157)	0.061	0.987	1.28		
		9	9	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>5</sub>	0.758 (0.703, 0.812)	5.77	8	0.6	0.982 (0.869, 1.095)	0.061	0.844	15.55		
		9	9	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>6</sub>	0.747 (0.693, 0.801)	2.22	8	0.5	0.992 (0.913, 1.071)	0.062	0.942	5.71		
		11	9	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>0</sub>	0.760 (0.705, 0.816)	12.51	8	0.1	1.042 (0.872, 1.212)	0.045	0.989	1.09		
		9	9	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>1</sub>	–	–	–	–	–	–	–	–		
		9	9	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>2</sub>	0.760 (0.705, 0.816)	12.51	8	0.1	1.042 (0.872, 1.212)	0.045	0.989	1.09		
		9	9	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>3</sub>	0.795 (0.742, 0.847)	2.51	8	0.9	0.981 (0.890, 1.071)	0.042	0.992	0.76		
Elley et al. [11]	0.914 (0.881–0.937)	10	10	FPG <sup>†</sup>	M <sub>4</sub>	0.806 (0.755, 0.857)	6.65	8	0.5	1.006 (0.865, 1.147)	0.043	0.992	0.73		
		9	9	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>5</sub>	0.820 (0.733, 0.867)	4.49	8	0.8	0.981 (0.868, 1.094)	0.041	0.902	9.70		
		9	9	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>6</sub>	0.789 (0.737, 0.842)	15.26	8	0.05	1.000 (0.895, 1.104)	0.042	0.963	3.61		
		12	10	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>0</sub>	0.744 (0.701, 0.788)	7.47	8	0.4	0.983 (0.895, 1.070)	0.049	0.994	0.58		
		10	10	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>1</sub>	0.744 (0.701, 0.788)	7.46	8	0.4	0.981 (0.895, 1.068)	0.049	0.994	0.58		
		10	10	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>2</sub>	0.744 (0.701, 0.788)	7.46	8	0.4	0.981 (0.895, 1.068)	0.049	0.994	0.58		
		10	10	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>3</sub>	0.752 (0.709, 0.793)	9.86	8	0.2	0.970 (0.890, 1.051)	0.049	0.994	0.55		
		13	13	BMI <sup>†</sup> , FPG <sup>†</sup> , Oral Diabetic drugs <sup>†</sup>	M <sub>4</sub>	0.774 (0.734, 0.814)	14.25	8	0.07	0.950 (0.838, 1.061)	0.051	0.994	0.53		
		10	10	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>5</sub>	0.762 (0.723, 0.801)	8.76	8	0.3	0.943 (0.812, 1.073)	0.049	0.942	5.73		
		10	10	UACR <sup>‡</sup> , HbA1c <sup>‡</sup>	M <sub>6</sub>	0.761 (0.722, 0.801)	10.05	8	0.2	0.976 (0.896, 1.056)	0.050	0.982	1.70		

**Table 3** (continued)

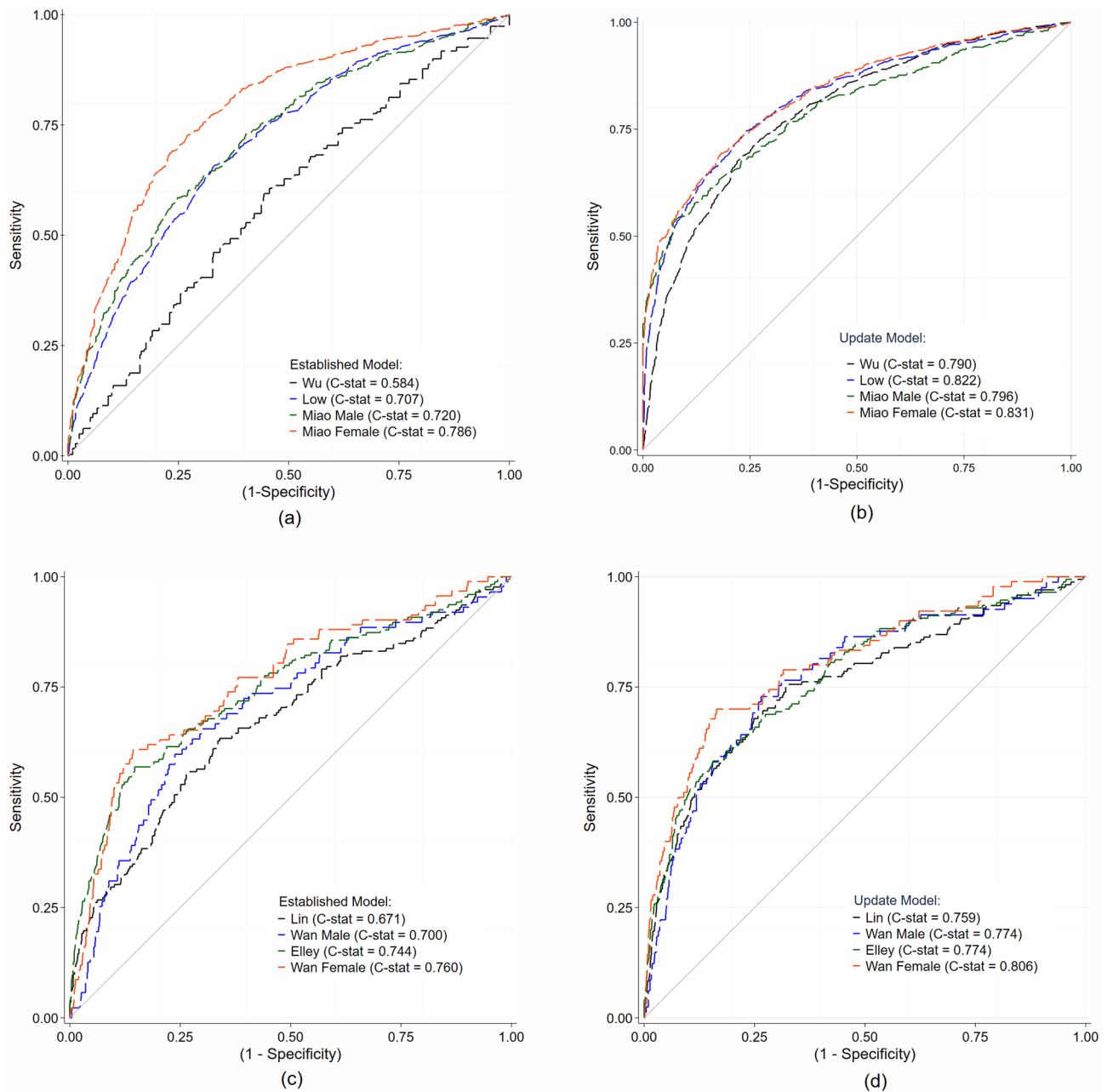
Study	Derivative	Prognostic factors		External validation								
		$n_1$	$n_2$	Method	C-statistics (95% CI)	Goodness of fit	O/E ratio (95% CI)	Brier Score	$S_{VH}$	Penalty (%)		
C-statistics (95% CI)		Variables missing/added in the equation		H–L df P								
<b>Part B: simplified prognostic models</b>												
<b>Chronic kidney disease (CKD)</b>												
Wu et al [10]	0.720 (0.696–0.744)	4	–	$M_0$	0.585 (0.565, 0.605)	7.72	8	0.4	1.000 (0.976, 1.025)	0.241	0.985	1.47
		4	–	$M_1$	–	–	–	–	–	–	–	–
		4	–	$M_2$	0.585 (0.565, 0.605)	7.72	8	0.4	1.000 (0.976, 1.025)	0.241	0.985	1.47
		4	–	$M_3$	0.603 (0.583, 0.623)	7.58	8	0.4	0.999 (0.975, 1.024)	0.239	0.989	1.00
		6	FPG * oral diabetic drugs <sup>†</sup>	$M_4$	0.663 (0.643, 0.682)	8.32	8	0.4	1.000 (0.969, 1.030)	0.227	0.987	1.21
		4	–	$M_5$	0.638 (0.619, 0.657)	2.88	8	0.9	1.000 (0.986, 1.014)	0.231	0.953	4.65
		4	–	$M_6$	0.633 (0.613, 0.652)	3.94	8	0.8	1.000 (0.971, 1.032)	0.232	0.972	2.70
<b>End Stage renal Disease (ESRD)</b>												
Lin et al. [12]	0.920 (0.900–0.930)	11	HbA1c <sup>‡</sup> , SBP (SD) <sup>‡</sup>	$M_0$	0.657 (0.610, 0.703)	5.13	8	0.7	1.010 (0.934, 1.086)	0.052	0.987	1.28
		9	HbA1c <sup>‡</sup> , SBP (SD) <sup>‡</sup>	$M_1$	–	–	–	–	–	–	–	–
		9	HbA1c <sup>‡</sup> , SBP (SD) <sup>‡</sup>	$M_2$	0.657 (0.610, 0.703)	5.13	8	0.7	1.010 (0.934, 1.086)	0.052	0.987	1.28
		9	HbA1c <sup>‡</sup> , SBP (SD) <sup>‡</sup>	$M_3$	0.745 (0.703, 0.788)	6.24	8	0.6	0.982 (0.902, 1.062)	0.051	0.993	0.66
		11	FPG <sup>†</sup> , BMI <sup>†</sup>	$M_4$	0.756 (0.713, 0.798)	12.99	8	0.1	0.971 (0.860, 1.082)	0.052	0.993	0.63
		9	HbA1c <sup>‡</sup> , SBP (SD) <sup>‡</sup>	$M_5$	0.774 (0.734, 0.812)	11.12	8	0.2	0.973 (0.863, 1.083)	0.048	0.876	12.31
		9	HbA1c <sup>‡</sup> , SBP (SD) <sup>‡</sup>	$M_6$	0.756 (0.714, 0.798)	5.78	8	0.6	0.976 (0.896, 1.056)	0.049	0.949	5.06

$n_1$  number of prognostic factors included in derivative,  $n_2$  number of prognostic factors included in external validation,  $S_{VH}$  Global heuristic shrinkage factor, M1 did not perform if the model did not report the baseline coefficients

<sup>‡</sup>Missing variables

<sup>†</sup>Updated by adding new variables into equation

\*Interaction effects



**Fig. 2** Receiving Operating Characteristic (ROC) curve comparisons between **a** baseline and **b** updated prognostic equations of CKD; and **c** baseline and **d** updated prognostic equations of ESRD

The coefficients for the associations between prognostic factors and CKD/ESRD in our cohort were estimated and compared to those in the original models, see Table S6. Our coefficients were mostly similar to the model proposed by Low and colleagues [8], but several predictors (i.e., sex, BMI, location, HDL-C, presence of hypertension, and/or dyslipidemia) were not significant compared to the models proposed by Miao et al. [9]. Most predictors in Wu's model were also significant in our data; however, the effect sizes

were lower for SBP, and diabetes duration and the direction of effect was reversed for BMI. Comparison of the corresponding rank odds ratio of predictors included in their respective CKD models identified creatinine ( $\beta = 4.653$ ) and retinopathy ( $\beta = 1.045$ ) with the strongest effects for females in Miao's model, whereas SBP ( $\beta = 0.902$ ) and diabetes duration ( $\beta = 0.891$ ) were highly associated with CKD in Wu's models, respectively (Table S6).

For modelling ESRD, only three of the 10 predictors were significant in Elley's [11] equations, including creatinine, diabetes duration and microalbuminuria, whereas in Wan's [13] models for female participants, insulin use, oral diabetic drug, and SBP were significantly correlated with ESRD in our multivariate analyses (Table S6).

## External validation

External validations were performed for models  $M_1$  to  $M_6$  where applicable (Table S7). Results of CKD-ESRD models are summarised in Table 3. At baseline ( $M_0$ ), all prognostic models showed fair calibration, but discrimination varied from poor to moderate, i.e., 0.585 to 0.707 and 0.671 to 0.760 for CKD and ESRD, respectively (Fig. 2). Sex-specific specific CKD and ESRD models performed better for females. For CKD, Miao's model for females generated a C-statistic of 0.786 (0.765–0.806) compared to 0.720 (0.691–0.749) for males, see Table 3.

All CKD-ESRD models provided improved C-statistics following additional adjustments of the regression coefficient ( $M_3$ ) and updated models from ( $M_4$ – $M_6$ ), see Figure S2. We updated CKD models by adding biomarkers (i.e., FPG groups  $< 126$  vs  $\geq 126$  mg/dL) and/or interaction effects with oral diabetic drug use; the greatest improvement was observed in the model by Wu and colleagues with a C-statistic of 0.790 (0.774–0.806), see Table 3.

In the baseline validation, most CKD models were well-calibrated in our population with O/E ranging from 0.999 (0.975–1.024) to 1.009 (0.929–1.090). Model calibration remained similar after updating, although Miao's model for males and females showed a slight overestimation of 1.052 (0.868–1.235), and 1.036 (0.917–1.156), respectively.

Four ESRD risk scores showed moderate to good calibration for baseline validation, recalibration, and updated models, see Figure S3 and Table 3. Fitting the equations using our validation set of ESRD equations ( $M_5$ ) showed worsening shrinkage, with a penalty of 12.31% and 15.55% for Lin's and Wan's male models, respectively.

The Brier score is another measure of prediction accuracy, ranging between 0 and 1, where lower scores indicate better accuracy. The Brier scores for the baseline and updated models are presented in Table 3. In the updated CKD model, the lowest Brier score was observed in Miao's model for females (0.162), Low's model (0.168), Miao's model for males (0.178), and Wu's model (0.185). Of the four ESRD models, the Brier score for the updated models ( $M_4$ ) was superior and ranged from 0.043 to 0.061.

Table S8 provides a summary of the model improvements implemented following baseline validation. New additional predictor variables (i.e., glucose level and/or interaction with oral diabetic medication) significantly improved the discrimination for the CKD models. The

highest improvement was observed in Wu's models with  $\Delta C$ -statistic of 0.214 (0.193–0.234). Most ESRD models showed minor significant discrimination improvements in the updated models.

## Discussion

We externally evaluated, validated, compared and updated six previously published models for predicting CKD/ESRD in a nationwide cohort of Thai participants with T2D, in line with recent framework guidelines [18, 19, 31, 38]. At baseline, most models provided only modest discrimination of T2D patients who developed CKD/ESRD. Two [10, 12] models demonstrated similar performance to their parent models. All models showed good calibration and upon modification, the agreement between observed and expected risk was fair, with only a few models showing slight overestimation.

In this study, the associations observed between prognostic factors and CKD/ESRD risk in Thai participants with T2D differed from previous studies. For instance, either hypertension or dyslipidaemia, LDL-C, and BMI were negatively associated with CKD risk in some models [8–10], with only a few predictors (i.e., diabetic duration, creatinine, and oral diabetic medications) significantly correlated with ESRD risk. We suspect that the lack of associations or variation in the direction of effect observed between previously reported predictor outcomes may have resulted from heterogeneity among the predictors and outcomes in our data, and that used previously for the development sets. However, we were unable to include two important biomarker predictor variables for four [8, 11–13] models (i.e., UACR and HbA1c) as they were unavailable in our data.

We postulate that the magnitude of the C-statistics and miscalibration observed may be explained by case-mix effects represented by the number of events, predictor effects, and heterogeneity in the population characteristics [19, 44, 45]. Variation of the included predictor variables, and sample size characteristics between derivation and validation settings, are likely responsible for the modest model performance in our population [19, 46].

In general, discrimination and calibration improved in our updated models. Although most models demonstrated lower discrimination in our data compared to their original settings, our updated models showed consistent improvement for all evaluation metrics (i.e., Brier score, shrinkage factor, penalty regression, and C-statistics). Most CKD-ESRD models also showed better reclassification (i.e.,  $\Delta C$ -statistic) for the enhanced models. Despite a lack of existing standards, Pencina et al. proposed that  $\Delta C$ -statistics greater than 0.01 represents a relevant improvement in model prediction [47, 48]. For our data, all models showed significant improvement on modification, with

$\Delta$ C-statistics ranging between 0.041 and 0.214 for CKD and 0.025 to 0.089 for ESRD equations.

The Brier score has been proposed as a measure of discrimination and calibration for model validation [49]. In this study, ESRD models performed better compared to those for CKD as determined by Brier scores. Almost every validation and updated model showed improved predictions (as judged by a Benchmark value less than 0.25) [40].

In our updated models, four proved more effective either for the prediction of CKD [8, 9] or ESRD [11, 13] in our population, without the need for recalibration or updated equations. These models consistently exceeded all others in terms of calibration and discrimination, and were more comparable to the derived models. Only Elley's model [11] provided a web calculator ([http://www.nzssd.org.nz/cvd\\_renal/](http://www.nzssd.org.nz/cvd_renal/)) to facilitate easier routine clinical practice use.

The strengths of our study include the long-term follow-up of diabetic progression in 26,170 individuals over 20 years, the definition of CKD from multiple data sources, and the evaluation of previously published prognostic models identified from a current SR/MA. This study was based on real world data from a clinical setting that used a broad range of routinely captured potential predictor variables evaluated for prognostic performance of renal outcomes in those with incident diabetes. To our knowledge, this is the first independent validation of CKD-ESRD prognostic models in an Asian population using real world data, beyond the populations from which the models originated. Therefore, our findings should be useful in predicting CKD-ESRD occurrence in other Asian regions where their settings are similar to Thailand.

Our study highlighted that eGFR assessment using creatinine was beneficial to kidney disease surveillance in a Thai population. By avoiding specific race/ethnicity coefficients, our updated models still offered accurate prognostic estimates which could be enhanced further through improved clinical and laboratory standards [30, 50].

Our study has several limitations. Markers of kidney damage, such as albuminuria and cystatin-C were not available in our data and missing data for some predictor variables precluded prognostic risk estimates for some models.

## Conclusions

In conclusion, we have provided an independent external validation of prognostic models for the prediction of incident CKD/ESRD in participants with T2D from Thailand. All evaluated prognostic models showed only moderate discriminative performance, but fair calibration at baseline validation. Updated prognostic scores improved predictive performance in most of the evaluation metrics (i.e., discrimination, calibration, and Brier score). An updated prognostic model for clinical use in Asian populations is provided.

Although no model was excellent, prognostic equations not delimited by sex (i.e., Low's [8] and Elley's [11]) performed better in our data and may offer clinical utility as a CKD screening tool in primary care for patients with diabetes.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s40620-021-01220-w>.

**Acknowledgements** This study was a part of a Ph.D. in the Data Science for Healthcare Program for the lead author (Sigit Ari Saputro) under the Department of Clinical Epidemiology and Biostatistics, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Thailand. The authors would like to thank Nattakrit Tongpoonsakdi and Stephen John Pinder for proofreading and conducting a comprehensive English language review.

**Author contributions** SAS, OP, and AT conceptualized the study design and take responsibility for the entire project. OP, AP, WA and AT prepared the data. SAS and AP performed data cleaning, aggregation and analysis. SAS drafted the manuscript with input from all authors. OP, AP, WA and AT edited and revised the manuscript. GM and JA edited and revised the manuscript for significant intellectual issues. All authors have read and approved the final paper.

**Funding** This research did not receive any specific grant funding from the public, commercial, or not-for-profit sectors.

## Declarations

**Conflict of interest** The authors have no conflicts of interest.

**Ethical approval** We explain in a protocol which already approved by the Human Research Ethics Committee of Ramathibodi Hospital (Number: #MURA2019/35).

**Informed consent** We use NHES as secondary data which they have individual consent from each participant and we explain in a protocol which already approved by the Human Research Ethics Committee of Ramathibodi Hospital (Number: #MURA2019/35).

**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/>.

## References

- George C, Mogueo A, Okpechi I, Echouffo-Tcheugui JB, Kengne AP (2017) Chronic kidney disease in low-income to middle-income countries: the case for increased screening. *BMJ Glob Health* 2(2):e000256. <https://doi.org/10.1136/bmjgh-2016-000256>

2. Hasan M, Sutradhar I, Gupta RD, Sarker M (2018) Prevalence of chronic kidney disease in South Asia: a systematic review. *BMC Nephrol* 19(1):291. <https://doi.org/10.1186/s12882-018-1072-5>
3. Jager KJ, Kovesdy C, Langham R, Rosenberg M, Jha V, Zoccali C (2019) A single number for advocacy and communication-worldwide more than 850 million individuals have kidney diseases. *Kidney Int* 96(5):1048–1050. <https://doi.org/10.1016/j.kint.2019.07.012>
4. International Diabetes Federation (2019) *IDF Diabetes Atlas, 9th edn*. International Diabetes Federation, Brussels, Belgium
5. Global Burden of Disease Collaborators (2017) Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global burden of disease study 2016. *Lancet* 390(10100):1151–1210. [https://doi.org/10.1016/s0140-6736\(17\)32152-9](https://doi.org/10.1016/s0140-6736(17)32152-9)
6. Basu S, Sussman JB, Berkowitz SA, Hayward RA, Bertoni AG, Correa A, Mwasongwe S, Yudkin JS (2018) Validation of risk equations for complications of type 2 diabetes (RECODE) using individual participant data from diverse longitudinal cohorts in the U.S. *Diabetes Care* 41(3):586–595. <https://doi.org/10.2337/dc17-2002>
7. Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ, McGinn T, Guyatt G (2017) Discrimination and calibration of clinical prediction models: users' guides to the medical literature. *JAMA* 318(14):1377–1384. <https://doi.org/10.1001/jama.2017.12126>
8. Low S, Lim SC, Zhang X, Zhou S, Yeoh LY, Liu YL, Tavintharan S, Sum CF (2017) Development and validation of a predictive model for chronic kidney disease progression in type 2 diabetes mellitus based on a 13-year study in Singapore. *Diabetes Res Clin Pract* 123:49–54. <https://doi.org/10.1016/j.diabres.2016.11.008>
9. Miao DD, Pan EC, Zhang Q, Sun ZM, Qin Y, Wu M (2017) Development and validation of a model for predicting diabetic nephropathy in Chinese people. *Biomed Environ Sci* 30(2):106–112. <https://doi.org/10.3967/bes2017.014>
10. Wu M, Lu J, Zhang L, Liu F, Chen S, Han Y, Zhao F, Guo K, Bao Y, Chen H, Jia W (2017) A non-laboratory-based risk score for predicting diabetic kidney disease in Chinese patients with type 2 diabetes. *Oncotarget* 8(60):102550–102558. <https://doi.org/10.18632/oncotarget.21684>
11. Elley CR, Robinson T, Moyes SA, Kenealy T, Collins J, Robinson E, Orr-Walker B, Drury PL (2013) Derivation and validation of a renal risk score for people with type 2 diabetes. *Diabetes Care* 36(10):3113–3120. <https://doi.org/10.2337/dc13-0190>
12. Lin CC, Li CI, Liu CS, Lin WY, Lin CH, Yang SY, Li TC (2017) Development and validation of a risk prediction model for end-stage renal disease in patients with type 2 diabetes. *Sci Rep* 7(1):10177. <https://doi.org/10.1038/s41598-017-09243-9>
13. Wan EYF, Fong DYT, Fung CSC, Yu EYT, Chin WY, Chan AKC, Lam CLK (2017) Prediction of new onset of end stage renal disease in Chinese patients with type 2 diabetes mellitus - a population-based retrospective cohort study. *BMC Nephrol* 18(1):257. <https://doi.org/10.1186/s12882-017-0671-x>
14. Hu Y, Shi R, Mo R, Hu F (2020) Nomogram for the prediction of diabetic nephropathy risk among patients with type 2 diabetes mellitus based on a questionnaire and biochemical indicators: a retrospective study. *Aging (Albany NY)* 12(11):10317–10336. <https://doi.org/10.18632/aging.103259>
15. Jiang S, Fang J, Yu T, Liu L, Zou G, Gao H, Zhuo L, Li W (2020) Novel model predicts diabetic nephropathy in type 2 diabetes. *Am J Nephrol* 51(2):130–138. <https://doi.org/10.1159/000505145>
16. Mogee A, Echouffo-Tcheugui JB, Matsha TE, Erasmus RT, Kengne AP (2015) Validation of two prediction models of undiagnosed chronic kidney disease in mixed-ancestry South Africans. *BMC Nephrol* 16:94. <https://doi.org/10.1186/s12882-015-0093-6>
17. Fraccaro P, van der Veer S, Brown B, Prosperi M, O'Donoghue D, Collins GS, Buchan I, Peek N (2016) An external validation of models to predict the onset of chronic kidney disease using population-based electronic health records from Salford, UK. *BMC Med* 14:104. <https://doi.org/10.1186/s12916-016-0650-2>
18. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, Woodward M (2012) Risk prediction models: II. External validation, model updating, and impact assessment. *Heart* 98(9):691–698. <https://doi.org/10.1136/heartjnl-2011-301247>
19. Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG (2015) A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J Clin Epidemiol* 68(3):279–289. <https://doi.org/10.1016/j.jclinepi.2014.06.018>
20. Collins GS, Omar O, Shanyinde M, Yu LM (2013) A systematic review finds prediction models for chronic kidney disease were poorly reported and often developed using inappropriate methods. *J Clin Epidemiol* 66(3):268–277. <https://doi.org/10.1016/j.jclin.2012.06.020>
21. Yang XL, So WY, Kong AP, Clarke P, Ho CS, Lam CW, Ng MH, Lyu RR, Yin DD, Chow CC, Cockram CS, Tong PC, Chan JC (2006) End-stage renal disease risk equations for Hong Kong Chinese patients with type 2 diabetes: Hong Kong diabetes registry. *Diabetologia* 49(10):2299–2308. <https://doi.org/10.1007/s00125-006-0376-3>
22. Jardine MJ, Hata J, Woodward M, Perkovic V, Ninomiya T, Arima H, Zoungas S, Cass A, Patel A, Marre M, Mancia G, Mogensen CE, Poulter N, Chalmers J (2012) Prediction of kidney-related outcomes in patients with type 2 diabetes. *Am J Kidney Dis* 60(5):770–778. <https://doi.org/10.1053/j.ajkd.2012.04.025>
23. Wysham CH, Gauthier-Loiselle M, Bailey RA, Manceur AM, Lefebvre P, Greenberg M, Duh MS, Young JB (2020) Development of risk models for major adverse chronic renal outcomes among patients with type 2 diabetes mellitus using insurance claims: a retrospective observational study. *Curr Med Res Opin* 36(2):219–227. <https://doi.org/10.1080/03007995.2019.1682981>
24. Dunkler D, Gao P, Lee SF, Heinze G, Clase CM, Tobe S, Teo KK, Gerstein H, Mann JF, Oberbauer R (2015) Risk prediction for early CKD in type 2 diabetes. *Clin J Am Soc Nephrol* 10(8):1371–1379. <https://doi.org/10.2215/cjn.10321014>
25. Nowak N, Skupien J, Smiles AM, Yamanouchi M, Niewczasz MA, Galecki AT, Duffin KL, Breyer MD, Pullen N, Bonventre JV, Krolewski AS (2018) Markers of early progressive renal decline in type 2 diabetes suggest different implications for etiological studies and prognostic tests development. *Kidney Int* 93(5):1198–1206. <https://doi.org/10.1016/j.kint.2017.11.024>
26. Peters KE, Davis WA, Ito J, Winfield K, Stoll T, Bringans SD, Lipscombe RJ, Davis TME (2017) Identification of novel circulating biomarkers predicting rapid decline in renal function in type 2 diabetes: the fremantle diabetes study phase II. *Diabetes Care* 40(11):1548–1555. <https://doi.org/10.2337/dc17-0911>
27. Basu S, Sussman JB, Berkowitz SA, Hayward RA, Yudkin JS (2017) Development and validation of risk equations for complications of type 2 diabetes (RECODE) using individual participant data from randomised trials. *Lancet Diabetes Endocrinol* 5(10):788–798. [https://doi.org/10.1016/s2213-8587\(17\)30221-8](https://doi.org/10.1016/s2213-8587(17)30221-8)
28. Keane WF, Zhang Z, Lyle PA, Cooper ME, de Zeeuw D, Grunfeld JP, Lash JP, McGill JB, Mitch WE, Remuzzi G, Shahinfar S, Snapinn SM, Toto R, Brenner BM (2006) Risk scores for predicting outcomes in patients with type 2 diabetes and nephropathy: the RENAAL study. *Clin J Am Soc Nephrol* 1(4):761–767. <https://doi.org/10.2215/cjn.01381005>
29. Liao LN, Li TC, Li CI, Liu CS, Lin WY, Lin CH, Yang CW, Chen CC, Chang CT, Yang YF, Liu YL, Kuo HL, Tsai FJ, Lin CC (2019) Genetic risk score for risk prediction of diabetic



- nephropathy in Han Chinese type 2 diabetes patients. *Sci Rep* 9(1):19897. <https://doi.org/10.1038/s41598-019-56400-3>
30. Delgado C, Baweja M, Burrows NR, Crews DC, Eneanya ND, Gadegbeku CA, Inker LA, Mendu ML, Miller WG, Moxey-Mims MM, Roberts GV, St Peter WL, Warfield C, Powe NR (2021) Reassessing the inclusion of race in diagnosing kidney diseases: an interim report from the NKF-ASN task force. *Am J Kidney Dis*. <https://doi.org/10.1053/j.ajkd.2021.03.008>
  31. Collins GS, Reitsma JB, Altman DG, Moons KG (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *J Clin Epidemiol* 68(2):134–143. <https://doi.org/10.1016/j.jclinepi.2014.11.010>
  32. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 162(1):W1-73. <https://doi.org/10.7326/M14-0698>
  33. Tantivess S, Yothisamut J, Saengsri W (2019) Utilisation of evidence from Thailand's national health examination survey in policy development: finding the weakest link. *Health Res Policy Syst* 17(1):104. <https://doi.org/10.1186/s12961-019-0512-4>
  34. Aekplakorn W, Chariyalertsak S, Kessomboon P, Assanangkornchai S, Taneepanichskul S, Putwatana P (2018) Prevalence of diabetes and relationship with socioeconomic status in the Thai population: national health examination survey, 2004–2014. *J Diabetes Res* 2018:1654530. <https://doi.org/10.1155/2018/1654530>
  35. Nguyen Ngoc H, Kriengsinyos W, Rojroongwasinkul N, Aekplakorn W (2019) Association of adiposity indices with hypertension in middle-aged and elderly Thai population: national health examination survey 2009 (NHES-IV). *J Cardiovasc Dev Dis*. <https://doi.org/10.3390/jcdd6010013>
  36. Lee CH, Shih AZL, Woo YC, Fong CHY, Yuen MMA, Chow WS, Lam KSL (2017) Which creatinine-based estimated glomerular filtration rate equation best predicts all-cause mortality in Chinese subjects with type 2 diabetes? *Diabetes Res Clin Pract* 126:25–29. <https://doi.org/10.1016/j.diabres.2017.01.010>
  37. Tomeleri CM, Ronque ER, Silva DR, Cardoso Júnior CG, Fernandes RA, Teixeira DC, Barbosa DS, Venturini D, Okino AM, Oliveira JA, Cyrino ES (2015) Prevalence of dyslipidemia in adolescents: comparison between definitions. *Rev Port Cardiol* 34(2):103–109. <https://doi.org/10.1016/j.repc.2014.08.020>
  38. Moons KG, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG, Grobbee DE (2012) Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart* 98(9):683–690. <https://doi.org/10.1136/heartjnl-2011-301246>
  39. Janssen KJ, Moons KG, Kalkman CJ, Grobbee DE, Vergouwe Y (2008) Updating methods improved the performance of a clinical prediction model in new patients. *J Clin Epidemiol* 61(1):76–86. <https://doi.org/10.1016/j.jclinepi.2007.04.018>
  40. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, Kattan MW (2010) Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 21(1):128–138. <https://doi.org/10.1097/EDE.0b013e3181c30fb2>
  41. Cook NR (2007) Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* 115(7):928–935. <https://doi.org/10.1161/circulationaha.106.672402>
  42. Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD (2004) Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. *Stat Med* 23(16):2567–2586. <https://doi.org/10.1002/sim.1844>
  43. Riley RD, Ensor J, Snell KIE, Harrell FE Jr, Martin GP, Reitsma JB, Moons KGM, Collins G, van Smeden M (2020) Calculating the sample size required for developing a clinical prediction model. *BMJ* 368:m441. <https://doi.org/10.1136/bmj.m441>
  44. Vergouwe Y, Moons KG, Steyerberg EW (2010) External validity of risk models: use of benchmark values to disentangle a case-mix effect from incorrect coefficients. *Am J Epidemiol* 172(8):971–980. <https://doi.org/10.1093/aje/kwq223>
  45. Candela-Toha A, Elías-Martín E, Abraira V, Tenorio MT, Parise D, de Pablo A, Centella T, Liaño F (2008) Predicting acute renal failure after cardiac surgery: external validation of two new clinical scores. *Clin J Am Soc Nephrol* 3(5):1260–1265. <https://doi.org/10.2215/cjn.00560208>
  46. Nieboer D, van der Ploeg T, Steyerberg EW (2016) Assessing discriminative performance at external validation of clinical prediction models. *PLoS One* 11(2):e0148820. <https://doi.org/10.1371/journal.pone.0148820>
  47. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS (2008) Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 27(2):157–172. <https://doi.org/10.1002/sim.2929> (**discussion 207–112**)
  48. Pencina MJ, D'Agostino RB, Pencina KM, Janssens AC, Greenland P (2012) Interpreting incremental value of markers added to risk prediction models. *Am J Epidemiol* 176(6):473–481. <https://doi.org/10.1093/aje/kws207>
  49. Hilden J, Gerds TA (2014) A note on the evaluation of novel biomarkers: do not rely on integrated discrimination improvement and net reclassification index. *Stat Med* 33(19):3405–3414. <https://doi.org/10.1002/sim.5804>
  50. Pottel H, Björk J, Courbebaisse M, Couzi L, Ebert N, Eriksen BO, Dalton RN, Dubourg L, Gaillard F, Garrouste C, Grubb A, Jacquemont L, Hansson M, Kamar N, Lamb EJ, Legendre C, Littmann K, Mariat C, Melsom T, Rostaing L, Rule AD, Schaeffner E, Sundin PO, Turner S, Bökenkamp A, Berg U, Åsling-Monemi K, Selistre L, Åkesson A, Larsson A, Nyman U, Delanaye P (2021) Development and validation of a modified full age spectrum creatinine-based equation to estimate glomerular filtration rate: a cross-sectional analysis of pooled data. *Ann Intern Med* 174(2):183–191. <https://doi.org/10.7326/m20-4366>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.