Supplementary material

Search strategy for PubMed via the National Library of Medicine

Coverage period: January 1, 2016 to February 1, 2026

Date last searched: February 1, 2026

Language limit: English

Records retrieved: 2,246

("Diabetes Mellitus, Type 2"[Mesh] OR "type 2 diabetes"[Title/Abstract] OR "T2DM"[Title/Abstract] OR "T2D"[Title/Abstract])

AND

("risk score"[Title/Abstract] OR "risk model"[Title/Abstract] OR "prediction model"[Title/Abstract] OR "predictive model"[Title/Abstract] OR "prognostic model"[Title/Abstract] OR "risk prediction"[Title/Abstract])

NOT

("systematic review"[Publication Type] OR "review"[Publication Type] OR "meta-analysis"[Publication Type] OR "editorial"[Publication Type] OR "letter"[Publication Type] OR "comment"[Publication Type] OR "case reports"[Publication Type] OR "clinical trial"[Publication Type] OR "randomized controlled trial"[Publication Type] OR "diabetes complications"[Title] OR "diabetic complications"[Title] OR "cardiovascular disease"[Title] OR "nephropathy"[Title] OR "diabetic kidney disease"[Title] OR "retinopathy"[Title] OR "neuropathy"[Title] OR "diabetic foot"[Title] OR "foot ulcer"[Title] OR "diabetic patients"[Title] OR "patients with diabetes"[Title] OR "known diabetes"[Title] OR "established diabetes"[Title] OR "diabetic population"[Title] OR "diabetic subjects"[Title])

AND

English[Language]

Filters: from 2016 - 2026

Search strategy for Embase via Embase.com

Coverage period: January 1, 2016 to February 1, 2026

Date last searched: February 1, 2026

Language limit: English

Records retrieved: 3,610

#1 'type 2 diabetes'/exp

#2 'type 2 diabetes':ti,ab OR 't2dm':ti,ab OR 't2d':ti,ab

#3 #1 OR #2

#4 'risk score':ti,ab OR 'risk model':ti,ab OR 'prediction model':ti,ab OR 'predictive model':ti,ab OR 'prognostic model':ti,ab OR 'risk prediction':ti,ab

#5 #3 AND #4

#6 'incident':ti,ab OR 'new onset':ti,ab OR 'risk of developing':ti,ab OR 'development of':ti,ab OR 'prediction of':ti,ab OR 'risk for':ti,ab

#7 #5 AND #6

#8 'systematic review':it OR 'review':it OR 'meta-analysis':it OR 'editorial':it OR 'letter':it OR 'comment':it OR 'case report':it OR 'clinical trial':it OR 'randomized controlled trial':it

#9 'diabetes complications':ti OR 'diabetic complications':ti OR 'cardiovascular disease':ti OR 'nephropathy':ti OR 'diabetic kidney disease':ti OR 'retinopathy':ti OR 'neuropathy':ti OR 'diabetic foot':ti OR 'foot ulcer':ti

#10 'diabetic patients':ti OR 'patients with diabetes':ti OR 'known diabetes':ti OR 'established diabetes':ti OR 'diabetic population':ti OR 'diabetic subjects':ti

#11 #7 NOT (#8 OR #9 OR #10)

#12 #11 AND [english]/lim AND [2016-2026]/py

Search strategy for Web of Science Core Collection via Clarivate

Coverage period: January 1, 2016 to February 1, 2026

Date last searched: February 1, 2026

Document type: Article

Language limit: English

Records retrieved: 2,780

(TS=("type 2 diabetes" OR "T2DM" OR "T2D"))

AND

(TS=("risk score" OR "risk model" OR "prediction model" OR "predictive model" OR "prognostic model" OR "risk prediction"))

NOT

(DT=(Review OR Editorial Material OR Letter OR Meeting Abstract)

OR

TI=("diabetes complications" OR "diabetic complications" OR "cardiovascular disease" OR "nephropathy" OR "diabetic kidney disease" OR "retinopathy" OR "neuropathy" OR "diabetic foot" OR "foot ulcer")

OR

TI=("diabetic patients" OR "patients with diabetes" OR "known diabetes" OR "established diabetes" OR "diabetic population" OR "diabetic subjects"))

AND

DT=(Article)

AND

LA=(English)

AND

PY=(2016-2026)

Supplementary Table1. Risk of Bias and Applicability Assessment of Included Studies Using the PROBAST Tool

Author (year)	Participants	Predictors	Outcome	Analysis	Overall Bias	Overall Applicability	Key concerns
Bernabe-Ortiz A (2016)	L	L	L	H	H	High	Stepwise variable selection; categorization of continuous predictors; limited calibration assessment; modest sample size for number of events.
Liu X (2016)	L	L	L	H	H	High	Stepwise variable selection; categorization of continuous predictors; lack of external validation; modest sample size for number of events.
Zhang M (2016)	L	L	L	H	H	High	Stepwise variable selection; no external validation in independent cohort; limited calibration assessment; FPG inclusion may

							inflate performance but is clinically relevant.
Han T (2016)	L	L	L	H	H	High	Cross-sectional derivation (prevalent cases); backward stepwise selection; potential overfitting; lack of temporality in derivation.
Miyakoshi T (2016)	L	L	L	H	H	High	Stepwise selection; dichotomization of continuous predictors; no external validation; limited calibration reporting.
Zhang H (2017)	L	L	L	H	H	High	Forward stepwise selection; categorization of continuous predictors; no external validation in independent cohort; limited calibration assessment.

Asadollahi K (2017)	H	L	L	H	H	High	Cross-sectional design (reverse causality); univariate pre-screening; categorization of continuous variables; unconventional score allocation (based on RR rather than coefficients); lack of calibration assessment.
Hippisley-Cox J (2017)	L	L	L	L	L	High	Exemplary methodology and validation.
Chen X (2017)	L	L	L	H	H	High	Categorization of continuous predictors; arbitrary score allocation based on β ranges; sub-cohort validation not fully independent; high loss to follow-up in sub-cohort.

Tian X (2017)	H	L	L	H	H	Low	Cross-sectional design; selection bias (enriched high-risk sample); categorization of continuous predictors; no internal/external validation; limited to oil field workers.
Yatsuya H (2018)	L	L	L	H	H	High	Backward stepwise selection; categorization of continuous predictors; no external validation; limited to Japanese male workers.
Abdullah N (2018)	H	L	L	H	H	High	Case-control design; no simplified risk score developed; missing data in physical activity; no validation; underpowered for gene-environment interactions.
Wang K (2018)	L	L	L	H	H	High	Stepwise variable selection; dichotomization of age; no external

							validation.
Han X (2018)	L	L	L	H	H	High	Categorization of continuous predictors; no external validation; no calibration assessment; age and waist circumference not retained despite univariate significance.
Sulaiman N (2018)	H	L	L	H	H	High	Cross-sectional design (reverse causality); stepwise selection; dichotomization of continuous variables; no calibration assessment; detects prevalent rather than incident diabetes.
Hu H (2018)	L	L	L	H	H	High	Backward stepwise selection; categorization of continuous predictors; no external validation.
Félix-Martínez GJ (2018)	H	L	L	L	H	High	Cross-sectional design; outcome based on single glucose

							measurement; potential selection bias due to survey non-response; self-reported variables (family history, blood pressure).
Kim SH (2018)	L	L	L	H	H	High	Categorization of continuous variables; stepwise selection; no external validation; no confidence intervals for c-index
Zhang M (2018)	H	L	L	H	H	High	Cross-sectional design (prevalent outcome); categorization of continuous predictors; no internal or external validation; no calibration assessment; EPV borderline
Wang K (2019)	L	L	L	H	H	High	Stepwise selection; no formal calibration assessment; temporal validation not fully independent

Kraege V (2019)	L	L	L	H	H	High	Variable categorization; no internal validation; no calibration assessment; external validation performed but not on all cohorts equally
Xie Z (2019)	H	L	L	H	H	High	Cross-sectional design; no external validation; missing data excluded; no calibration assessment; synthetic oversampling may introduce bias
Pei D (2019)	H	L	L	H	H	High	Cross-sectional design; missing data excluded; no external validation; data-driven tree pruning may overfit
Wang Y (2019)	L	L	L	H	H	High	Nested case-control design; no external validation; continuous variables categorized; no

							calibration assessment
Wang J (2019)	H	L	L	H	H	High	Cross-sectional design (prevalent outcome); stepwise selection; low AUC (0.65-0.67); no external validation
Arellano-Campos O (2019)	L	L	L	H	H	High	Categorization of continuous predictors; arbitrary point assignment; no external validation; no calibration assessment
Cai X (2020)	L	L	L	H	H	High	No external validation; retrospective design with potential selection bias; modest number of events
Bahijri S (2020)	H	L	L	H	H	High	Cross-sectional design; categorization of continuous predictors; arbitrary point assignment; no internal or external

							validation; no calibration assessment
Hu H (2020)	H	L	L	H	H	High	Highly selected "low risk" population; categorization of continuous predictors; no internal or external validation; small number of events
Liu X (2020)	L	L	L	H	H	High	Stepwise selection; categorization of continuous predictors; no external validation
Wilkinson L (2020)	L	L	L	L	L	High	Robust methodology with external validation and continuous predictors
Zhang L (2020)	H	L	L	H	H	High	Cross-sectional design; SMOTE may introduce bias; no external validation; no calibration assessment; black-box models

Shao X (2020)	L	L	L	H	H	High	Categorization of continuous predictors; multiple models without correction for multiple comparisons; variable selection based on different criteria across models
Wang H (2020)	L	L	L	H	H	High	Univariate pre-screening for variable selection; categorization of FBG and HbA1c; no external validation; no calibration assessment
Yang T (2020)	H	L	L	H	H	High	Cross-sectional design; feature selection using univariate pre-screening; complete case analysis may introduce bias; no calibration assessment

Del Razo-Olvera FM (2020)	L	L	L	H	H	High	Stepwise selection; TG categorized; short follow-up period (2.4 years)
Katsimpris A (2021)	H	L	L	H	H	High	Cross-sectional design; no external validation; no calibration assessment; model excludes established T2D risk factors
Rhee SY (2021)	L	L	L	H	H	High	No external validation; black-box nature of deep learning (partially addressed by LRP); no simplified clinical score
Awad SF (2021)	H	L	L	H	H	High	Simulated data for model development; no calibration assessment; validation data partially overlapped with model calibration data
Cai X (2021)	L	L	L	H	H	High	No external validation; retrospective design with potential

							selection bias; modest number of events
Wang Y (2021)	H	L	L	H	H	High	Cross-sectional design; no external validation; region-specific population may limit generalizability
Cai X (2021)	L	L	L	H	H	High	No external validation; retrospective design with potential selection bias; LASSO selection may be data-driven
Tavares DPS (2021)	H	L	L	H	H	High	Cross-sectional design (prevalent outcome); stepwise selection; no external validation; outcome based on administrative data not laboratory criteria
Mustafina SV (2021)	L	L	L	H	H	High	Variable selection based on univariate significance;

							dichotomization of continuous predictors using ROC-derived cut-offs (data-driven); no internal validation to correct for optimism; calibration not assessed.
Deberneh HM (2021)	H	L	H	H	H	Low	Outcome defined based on current-year FPG; essentially a cross-sectional classification rather than prediction; No external validation; class imbalance handled with SMOTE/undersampling; no calibration assessment; models not presented in usable format for clinical application.
Oh TJ (2021)	L	L	L	H	H	High	High loss to follow-up (35%); variables selected based on

							univariate significance; no calibration reported; no external validation; model discrimination modest (AUC 0.657).
Stiglic G (2021)	L	L	H	L	H	High	Outcome self-reported "diabetes or high blood sugar"; complete case analysis removed 35% participants may introduce selection bias; no external validation; LASSO selection data-driven; calibration plots show strong overprediction in uncalibrated models.
Al-Lawati JA (2021)	L	L	L	H	H	High	Cross-sectional design (prevalent outcome) cannot assess temporality; stepwise variable selection; 20.9% excluded due to missing outcome data may

							introduce selection bias; external validation showed performance decline (AUC 0.80 to 0.74).
Sulistiowati E (2021)	L	L	L	H	H	High	Bivariate screening then stepwise selection; no internal validation; calibration not reported; low specificity (50.3%)
Tan C (2022)	L	L	L	H	H	High	No external validation; validation using split-sample from same cohort may overestimate performance; LASSO selection data-driven but with 10-fold CV; calibration reported but Hosmer-Lemeshow test underpowered; decision curve analysis performed.
Wang Y (2022)	L	L	L	L	L	High	Well-conducted prospective cohort

							with large sample; calibration not reported; 38% loss to follow-up; external validation needed
Nicolaisen SK (2022)	L	L	L	L	L	High	Well-conducted study with large population-based cohort; 38% survey response rate; external validation needed; no simplified clinical score.
Zhao Y (2022)	L	L	L	L	L	High	Well-conducted external validation of existing risk score in large prospective cohort; AUC modest (0.675).
Sadek K (2022)	H	L	L	L	H	High	Cross-sectional design; family history missing >20%; continuous variables dichotomized; external validation performed.

Edlitz Y (2022)	L	L	L	L	L	High	Healthy volunteer bias; external validation showed performance degradation (AUC 0.87 to 0.75); anthropometric model not externally validated.
Liu Q (2022)	L	L	L	H	H	High	Retrospective design with potential selection bias; no external validation; no simplified clinical tool; calibration not assessed.
Martínez-Hervás S (2022)	L	L	L	H	H	High	Well-conducted prospective cohort; single FPG measurement; 44% loss to follow-up; no external validation; no calibration assessment.
Dong W (2022)	L	L	L	H	H	High	Cross-sectional design; no external validation; single blood test may misclassify.

Tong C (2022)	L	L	L	H	H	High	No external validation; sCr quartiles data-driven; no calibration assessment.
Boersma HE (2022)	L	L	H	H	H	High	50% incident cases self-reported; exclusion of participants with missing follow-up; no external validation; data-driven SAF cut-offs.
Wang S (2023)	H	L	H	H	H	High	Cross-sectional design; case group includes previously diagnosed T2DM (recall bias, treatment effects); SMOTE for imbalance; no external validation; single FPG measurement for diagnosis.
Zheng M (2023)	L	L	L	L	L	High	Well-conducted development and validation with large sample, external validation, and

							appropriate methods.
Yonel Z (2023)	L	L	L	L	L	High	Well-conducted development with multiple imputation, shrinkage, and external validation
Hu Y (2023)	H	L	L	H	H	High	Retrospective design; high missing rates for some variables (smoking 72%, drinking 72%, AST 58%, HDL-c 45%) addressed with multiple imputation; no external validation; sensitivity/specificity moderate (65–69%).
Yang J (2024)	L	L	L	H	H	High	No external validation; missing data handled with regression imputation; single-center study may limit generalizability.
Ojurongbe TA (2024)	H	L	H	H	H	Low	Case-control design; single hospital in Nigeria limits

							generalizability; small sample size with extremely high AUC suggests overfitting.
Ghiasi Hafezi S (2024)	L	L	L	H	H	High	No external validation; single-center study (Mashhad) may limit generalizability; very high specificity (99-100%) suggests potential overfitting; no calibration assessment.
Xu J (2024)	L	L	L	L	L	High	Well-conducted development with multiple imputation for missing data, bootstrap optimism correction, internal-external cross-validation, and external validation in independent cohort.
Lin Y (2024)	H	L	L	H	H	High	Retrospective design with potential selection bias; no external

							validation; single-center study limits generalizability.
Xie R (2025)	L	L	L	L	L	Low	Well-conducted large-scale study with internal and external validation, robust variable selection via LASSO with bootstrap, and comprehensive performance evaluation.
Kawase S (2024)	L	L	L	H	H	High	No external validation; single-center study limits generalizability; no sensitivity analysis for missing data; retrospective design; FBG/HbA1c measured only once at baseline.
Wang Y (2025)	L	L	L	H	H	High	No external validation; single-center data; outcome based on single FPG/HbA1c

							measurement and self-report; low event rate (1.7%) may limit power.
Shahraki PK (2025)	L	L	L	H	H	Low	No external validation; high loss to follow-up (31.8%); single-center study; variables categorized based on optimal cutoffs (data-driven).
Chee HK (2025)	L	L	L	H	H	High	Cross-sectional design; no external validation; single-center study ; small number of confirmed diabetes cases; outcome confirmation based on point-of-care HbA1c and medical referral compliance.
Munekawa C (2025)	L	L	L	L	L	High	Well-conducted large-scale study with internal and external validation; robust statistical methods; comprehensive variable

							selection with bootstrap resampling.
Kiviniemi E (2025)	L	L	L	L	L	High	Well-conducted study with large sample size, multiple external validation cohorts, robust variable selection, and comprehensive performance evaluation.
Alareeki A (2025)	H	L	L	H	H	Low	Model derived from simulated data; no external validation.
Chen YE (2025)	L	L	L	L	L	High	Well-conducted development and validation with large sample sizes, multiple external validations, and rigorous methodology.
Han X (2025)	L	L	L	L	L	High	Well-conducted study with robust meta-analysis, external validation in three diverse Chinese populations, and comprehensive

							comparison with existing model and ML methods.
Zhang X (2025)	L	L	H	H	H	High	Outcome based solely on single FPG measurement; no external validation; small sample size relative to events (107 events).
AlKetbi LM (2025)	L	L	H	H	H	Low	Outcome based on HbA1C >6.4% or documented diagnosis; no external validation; retrospective design; reliance on electronic health.
Satoh T (2025)	L	L	L	L	L	High	Well-conducted large-scale study with robust methodology, including Cox regression accounting for competing risk of death, internal validation, and comprehensive predictor selection.

L: low risk of bias; H: high risk of bias; N: unclear risk of bias.

Abbreviations: PROBAST, Prediction model Risk Of Bias ASsessment Tool; L, low risk of bias; H, high risk of bias; N, unclear risk of bias; T2DM, type 2 diabetes mellitus; AUC, area under the curve; EPV, events per variable; SMOTE, synthetic minority over-sampling technique; LASSO, Least Absolute Shrinkage and Selection Operator; CV, cross-validation; ROC, receiver operating characteristic; SAF, skin autofluorescence; FPG, fasting plasma glucose; HbA1c, glycated hemoglobin.

Supplementary Table2 Predictor Variables, Model Presentation, and Clinical Intended Use of the Included Models

First Author	Predictor Variables	Modifiable Predictors	Model Presentation	Intended Use
Bernabe-Ortiz A (2016)	Age (≥55 years vs. <55 years), diabetes in first-degree relatives (yes/no, including parents and/or siblings), WC (<90 cm, 90–99.9 cm, ≥100 cm)	WC	Points-based scoring system: age ≥55 years = 1 point, diabetes in relatives = 1 point, waist 90–99.9 cm = 1 point, waist ≥100 cm = 2 points	To screen for undiagnosed T2D in resource-constrained settings without requiring blood tests; also predicts incident diabetes over 2.5 years of follow-up.
Liu X (2016)	Age (55–65 years, 66–75 years, ≥76 years), BMI (<18.0 kg/m² [thin], 18.0–23.9 kg/m² [normal], 24.0–27.9 kg/m² [overweight], ≥28.0 kg/m² [obese]), FPG (normal vs. impaired [6.1–7.0 mmol/L]), self-rated health (healthy vs. unhealthy), physical activity (frequently vs. not frequently)	BMI, fasting plasma glucose, physical activity	Points-based scoring system: age 55–65 = 15 points, age 66–75 = 7 points; BMI 24.0–27.9 = 8 points, BMI ≥28.0 = 7 points, BMI <18.0 = –4 points; impaired fasting glucose = 7 points; unhealthy self-rated health = 6 points;	To predict 20-year risk of incident T2D in Chinese adults aged ≥55 years, accounting for the competing risk of non-diabetes death; to identify high-risk individuals (score ≥19) for targeted prevention and intervention.

			not frequently active = 3 points	
Zhang M (2016)	Age (years), BMI (kg/m²), TG (mmol/L), FPG(mmol/L)	BMI, TG, fasting plasma glucose	Risk score formula: score = 3×age + 12×BMI + 24×TG + 76×FPG (round each component to nearest integer)	To predict 6-year risk of incident T2D in rural Chinese adults; to identify high-risk individuals (score >936) ; suitable for primary care settings.
Han T (2016)	Classic noninvasive factors: age (years), gender (male/female), hypertension (self-report, BP ≥140/90 mmHg, or medication), family history of diabetes (first- or second-degree relatives), alcohol consumption (current drinker: ≥1	BMI, WC, alcohol consumption, regular exercise, hypertensi	Points-based scoring system (β coefficients ×10 for continuous variables; integer points for binary/categorical); total score calculated as sum of	To predict prevalent (derivation) and incident (validation) T2D risk in Northern Chinese adults; to identify high-risk individuals for targeted lifestyle

	drink/month in past year), regular exercise (≥30 minutes, ≥3 days/week), BMI(kg/m²), WC (cm). Dietary factors (liang/day, 1 liang ≈ 50 g): staple foods (white rice + wheaten food), livestock and its products, eggs, potato and its products, fresh fruit, fresh vegetable, dairy and its products (binary: any daily consumption vs. none).	on, and all dietary factors (staple foods, livestock, eggs, potato, fresh fruit, fresh vegetable, dairy)	points for each factor	interventions, particularly focusing on modifiable dietary and behavioral factors; designed for primary care and community screening settings.
Miyakoshi T (2016)	RSFull: Family history of diabetes (yes/no), age ≥55 years, SBP ≥130 mmHg, male sex, current smoking, FPG ≥5.3 mmol/L, 2hPG during 75g OGTT ≥7.6 mmol/L, HbA1c ≥5.6% (38 mmol/mol).	Smoking, SBP, FPG, 2hPG, HbA1c, TG, BMI	Points-based scoring system (predictors assigned 1–5 points based on HR magnitude); total score used to estimate 5-year diabetes	To predict incident T2D in Japanese adults; RSNI suitable for initial community screening; RS-2hPG for settings without OGTT; RSFull

	RS-2hPG: FH, age ≥55y, SBP ≥130 mmHg, male sex, FPG ≥5.3 mmol/L, HbA1c ≥5.6%, TG ≥1.13 mmol/L. RSNI: FH, age ≥55y, SBP ≥130 mmHg, male sex, BMI ≥25 kg/m².		probability	for most accurate prediction when OGTT is available.
Zhang H (2017)	Drinking tea frequently (≥3 times/week for >6 months), BMI (kg/m², ≥28.0 [obese]), WHtR (calculated as WC/height, ≥0.5), TG (mmol/L, 1.70–2.25, ≥2.26), FPG (mmol/L, 5.6–6.0, ≥6.1). Reference categories: no frequent tea drinking, BMI <28.0, WHtR <0.5, TG <1.70 mmol/L, FPG <5.6 mmol/L.	Drinking tea, BMI, WHtR, TG, FPG	Points-based scoring system: frequent tea drinking = –2, BMI ≥28.0 = 7, WHtR ≥0.5 = 7, TG 1.70–2.25 = 4, TG ≥2.26 = 6, FPG 5.6–6.0 = 11, FPG ≥6.1 = 25	To predict 6-year risk of incident T2D in rural Chinese adults; to identify high-risk individuals (score >16) for targeted lifestyle interventions.

Asadollahi K (2017)	Family history of diabetes (yes/no), place of life (urban/rural), age (years, ≤40, 41–50, 51–60, >60), hypertension (yes/no), daily exercise (≥30 min/day, yes/no), marital status (married/unmarried/widowed), gender (male/female), WC (cm by sex: male <95, 95–102, >102; female <80, 80–88, >88), smoking (yes/no), BMI (kg/m², <25, 25–30, >30).	BMI, WC, daily exercise, smoking, hypertension	Points-based scoring system; individual scores: family history = 3, age 41–50 = 2, age 51–60 = 2.5, age >60 = 3, hypertension = 3, no daily exercise = 1.5, male = 1, urban residence = 2, WC (moderate) = 1.5, WC (high) = 2.5, BMI 25–30 = 1, BMI >30 = 1.5, smoking = 1.5	To provide a simple, non-invasive self-assessment tool for the general population to estimate their current risk of having undiagnosed diabetes or IFG; to identify high-risk individuals who should seek confirmatory blood glucose testing.

Hippisley-Cox J (2017)	Model A: age (years), ethnicity (self-assigned: White/not recorded, Indian, Pakistani, Bangladeshi, other Asian, Caribbean, Black African, Chinese, other), deprivation (Townsend score), BMI (kg/m² with fractional polynomial terms), smoking status (non-smoker, former smoker, light [1–9/day], moderate [10–19/day], heavy [≥20/day]), family history of diabetes in first-degree relative (yes/no), cardiovascular disease (yes/no), treated hypertension (yes/no), regular use of corticosteroids (yes/no), prescription of atypical antipsychotics (yes/no), prescription of statins (yes/no), schizophrenia or bipolar affective disorder (diagnosis, yes/no), learning	BMI, smoking, fasting glucose, HbA1c (and the underlying conditions such as hypertension, which is modifiable through treatment)	Cox regression equations with baseline survival function; published as open source software; freely accessible web calculator at http://qdiabetes.org/2018	To estimate 10-year absolute risk of T2D in adults aged 25–84 for risk communication and clinical decision-making; Model A is intended for initial risk stratification to identify those needing blood glucose testing; Models B and C provide refined risk estimates after blood test results are available, enabling personalised preventive interventions and more accurate targeting of high-risk individuals.

	disability (diagnosis, yes/no), gestational diabetes (women only, yes/no), polycystic ovary syndrome (women only, yes/no).			
Chen X (2017)	Model B: includes all Model A variables plus FBG (mmol/L). Model C: includes all Model A variables plus HbA1c (mmol/mol). Non-invasive model: age (≥40 years vs. <40), BMI (<25 [reference], 25.0–29.9 [overweight], ≥30 [obese] kg/m²), family history of T2DM (yes/no in first-degree relatives), meat diet (meat mainly vs.	BMI, meat diet, hypertension, FPG	Points-based scoring system: age ≥40 = 4, overweight = 2, obesity = 4, family history = 3, meat diet = 3, hypertension = 2, IFG = 4	To predict incident T2D risk in rural Chinese adults; non-invasive model for initial screening in resource-limited primary care settings;

	vegetable mainly/equally), hypertension (SBP ≥140 mmHg, DBP ≥90 mmHg, or antihypertensive medication).			
	Plus-FPG model: all above + impaired fasting glucose (IFG, FPG 6.1–6.9 mmol/L).			
Tian X (2017)	Sex (female vs. male), age (<45, 45–54, 55–64 years), BMI (<25, ≥25 to <30, ≥30 kg/m²), history of high blood glucose (HHBG, yes/no), smoking (yes/no, current or former vs. never), TG (≥1.70 vs. <1.70 mmol/L), FPG (≥5.6 vs. <5.6 mmol/L).	BMI, smoking, TG, FPG	Points-based scoring system: sex female = 3, age 45–54 = 3, age 55–64 = 10, HHBG = 6, smoking = 4, BMI ≥25 to <30 = 2, BMI ≥30 = 6, TG ≥1.70 = 7, FPG ≥5.6 = 19	plus-FPG model for more accurate risk assessment when blood glucose testing is available; to identify high-risk individuals (score ≥6 for non-invasive, ≥7 for plus-FPG) for targeted lifestyle interventions and confirmatory testing. To screen for dysglycemia (pre-diabetes or undiagnosed diabetes) in oil field workers; to identify high-risk individuals for confirmatory OGTT and early lifestyle or pharmacological intervention.

Yatsuya H (2018)	Age (years), BMI (<25.0, 25.0–27.5, ≥27.5 kg/m²), smoking status (non-current/current), family history of diabetes in first-degree relatives (yes/no), TG (<150, ≥150 mg/dL), FBG (<100, 100-110, 110-126 mg/dL).	BMI, smoking, TG, FBG	Points-based scoring system: points assigned based on β coefficients divided by base regression unit (0.076 per 5-year age increase)	To estimate 10-year risk of incident T2D in middle-aged Japanese male workers; to identify high-risk individuals (score ≥30 corresponding to >20% risk); suitable for use in occupational health settings with annual health check-ups.
Abdullah N (2018)	Environmental risk factors: gender (male/female), age (years), WHR (continuous), physical inactivity (<150 min/week of vigorous activity vs. ≥150 min/week). Genetic risk factor: weighted genetic risk score (GRS) derived from 62 T2D-associated SNPs.	WHR, physical inactivity	Logistic regression coefficients; GRS calculated as weighted sum of risk alleles; no simplified points-based score provided	To assess the relative contribution of genetic and environmental factors to T2D risk in multi-ethnic Malaysian populations; to investigate gene-environment interactions; findings suggest limited clinical utility of genetic scores beyond

				established environmental risk factors for risk prediction.
Wang K (2018)	Males: age (≥50 years vs. <50), hypertension (SBP ≥140 mmHg, DBP ≥90 mmHg, or antihypertensive medication), 2hPG ≥7.8 mmol/L, FBG ≥6.1 mmol/L, TG ≥1.7 mmol/L. Females: age (≥55 years vs. <55), FBG ≥6.1 mmol/L, 2hPG ≥7.8 mmol/L, TG ≥1.7 mmol/L.	Hypertension, FBG, 2hPG, TG	Nomogram; points-based scoring system (separate for males and females) with total score ranges and corresponding 3-year T2DM risk categories (<10%, 10–20%, 20–30%, 30–40%, 40–50%, >50%)	To estimate 3-year risk of incident T2D in Chinese adults using readily available clinical measures; particularly useful in health check-up settings where OGTT is performed.
Han X (2018)	BMI (<24.0, 24.0–27.9 [overweight], ≥28.0 [obese] kg/m²), FBG (<5.6, 5.6–6.0, 6.0–6.9 mmol/L), hypertension (self-report, BP ≥140/90 mmHg, or antihypertensive medication), hyperlipidemia (TC >5.72 mmol/L, TG >1.70 mmol/L, lipid-lowering	BMI, FBG, hypertension, hyperlipidemia, smoking	Points-based scoring system: BMI 24.0–27.9 = 5, BMI ≥28.0 = 7; FBG 5.6–6.0 = 9, FBG 6.0–6.9 = 19; hypertension = 1; hyperlipidemia = 2; current smoking = 2; family history = 5	To estimate 5-year risk of incident T2D in middle-aged and older Chinese adults; to identify high-risk individuals (score ≥15); designed for primary care and community screening settings.

	medication, or physician diagnosis), current smoking (yes vs. no), family history of diabetes (yes/no in first-degree relatives).			
Sulaiman N (2018)	Age (≥35 vs. <35 years), family history of diabetes in parents or siblings (yes/no), hypertension (self-report and/or BP ≥140/90 mmHg), BMI (≥30 vs. <30 kg/m²), WHR (≥0.90 for males / ≥0.85 for females vs. below these cut-offs).	BMI, WHR, hypertension	Points-based scoring system: age ≥35 = 4, family history = 7, hypertension = 4, BMI ≥30 = 7, elevated WHR = 5; total score range 0–27	To screen for undiagnosed T2D among UAE citizens using simple, non-invasive measures; cut-off ≥19 indicates high risk of undiagnosed diabetes.
Hu H (2018)	Non-invasive model: sex, age (30–39, 40–49, 50–59 years), BMI (<21, 21–<23, 23–<25, 25–<27, 27–<29, ≥29 kg/m²), abdominal obesity (WC ≥90 cm for men, ≥80 cm for women), smoking status (never, former, current), hypertension status (yes/no).	BMI, abdominal obesity, smoking, hypertension, dyslipidem	Cox proportional hazards regression equations	To predict 7-year risk of incident T2D in the Japanese working population; non-invasive model for initial screening; invasive models for more accurate risk stratification.

	Invasive models: all above + dyslipidemia (yes/no), FPG (<100, 100–<110, 110–<126 mg/dL), HbA1c (<5.6, 5.6–<6.0, 6.0–<6.5%).	ia, FPG, HbA1c		
Félix-Martínez GJ (2018)	Model 1: age (years), WC (cm), SBP (SBP, mmHg). Model 2: age (years), WC (cm), height (cm), family history of diabetes (yes/no).	WC, SBP	Logistic regression equations with coefficients provided; interactive web applications available at http://gifelix.shinyapps.io/UDMScreeningMexico2006 and http://gifelix.shinyapps.io/UDMScreeningMexico2012	To screen for undiagnosed prevalent T2D in Mexican adults using simple, non-invasive measures; to identify individuals who should undergo confirmatory blood glucose testing; to provide a cost-effective tool for resource-limited settings.
Kim SH (2018)	Clinical model: age (years), sex (male/female), FBS (mg/dL, >80–85, >85–90, >90), ALT (IU/L, >30–40, >40), γ-GTP (IU/L,	FBS, ALT, γ-GTP, HDL, TG, SBP, BMI	Cox regression equations; points-based scoring system (points assigned based on β coefficients; total points	To predict 5-, 7-, and 10-year risk of incident T2D in Korean adults; to identify high-risk individuals for targeted

	sex-specific categorical), HDL (mg/dL, sex-specific >40/50), TG (mg/dL, >90–110, >110–160, >160–230, >230), SBP (mmHg, >130–150, >150), BMI (kg/m², >23–25, >25–30, >30), parental history of DM (yes/no). Integrated model: all above + SNP point-based score (sum of risk alleles from 5 SNPs: rs10975266, rs9311835, rs11057302, rs17154562, rs8038172; range 0–10).		range –1 to 38 for clinical model, –1 to 46 for integrated model)	prevention; to evaluate the incremental value of genetic factors over clinical factors alone.
Zhang M (2018)	Age (≥45 vs. <45 years), family history of diabetes (yes/no), hypertension (yes/no), WC (WC, ≥85 cm men / ≥80 cm women), BMI (<24, ≥24 to <28, ≥28 kg/m²), current smoking (yes/no), daily vegetable intake (<400 g vs. ≥400 g), daily fruit intake	BMI, WC, smoking, vegetable intake, fruit intake, hypertensi	Points-based scoring system: age ≥45 = 11, family history = 9, hypertension = 5, current smoking = 8, WC elevated = 7, BMI ≥24 = 2, BMI ≥28 = 8, vegetable <400	To screen for undiagnosed prevalent T2D in young and middle-aged Chinese adults using simple, non-invasive, and self-assessable variables.

	(<300 g vs. ≥300 g).	on	g = 8, fruit <300 g = 6.	
Wang K (2019)	Females: age (years), BMI (kg/m²), FBG (mmol/L), TG (mmol/L). Males: age (years), BMI (kg/m²), FBG (mmol/L), HDL cholesterol (mmol/L), TG (mmol/L).	BMI, FBG, HDL, TG	Nomogram: each continuous variable corresponds to points	To predict 3-year risk of incident T2D in Central Chinese adults; to identify high-risk individuals for targeted lifestyle interventions
Kraege V (2019)	Age (35–44, 45–54, 55–64, 65–75 years), sex (male/female), WC (sex-specific: e.g., <80, 80–89, 90–99, 100–109, ≥110 cm for women; <90, 90–99, 100–109, 110–119, ≥120 cm for men), family history of diabetes (parents or siblings, yes/no), hypertension (SBP ≥130 mmHg or DBP ≥85 mmHg or antihypertensive medication), physical inactivity (<2	WC, hypertensi on, physical activity	Points-based scoring system: points assigned for each category	To screen for individuals at high risk of incident T2D using simple, non-invasive, and inexpensive clinical data; applicable in resource-limited settings and across diverse populations.

	sessions of ≥20 min leisure-time physical activity per week).			
Xie Z (2019)	Age (31–40, 41–50, 51–60, 61–70, 71–80, >80 years), BMI (underweight, normal, overweight, obese), annual household income (8 categories), marital status (married, divorced, widowed, separated, never married, unmarried couple), race/ethnicity (white, black, American Indian/Alaskan Native, Asian, Native Hawaiian/other Pacific Islander, other, multiracial, Hispanic), sleep duration (1–6, 7–8, ≥9 h/day), time since last checkup (<1 year, 1–2 years, 3–5 years, >5 years, never), physical activity	BMI, sleep duration, checkup frequency, physical activity, smoking	Multiple machine learning models (SVM, decision tree, logistic regression, random forest, neural network, Naïve Bayes); no simplified points-based score provided	To screen for prevalent T2D using easily obtainable self-reported survey data; to identify individuals at high risk for diabetes who may benefit from early intervention; to discover novel risk factors (sleep duration, checkup frequency) for further investigation.

	(yes/no), smoking status (current daily, current some days, former, never), and other self-reported variables.			
Pei D (2019)	Age (≤34, 35–49, >49 years), gender (male/female), BMI (<25 vs. ≥25 kg/m²), hypertension (SBP ≥140 mmHg or DBP ≥90 mmHg or antihypertensive medication), history of cardiovascular disease or stroke (yes/no), family history of diabetes (yes/no), physical activity (≥30 min exercise ≥3 days/week vs. less), work-related stress (low, moderate, high based on self-report), salty food preference (yes/no).	BMI, hypertension, physical activity, work-related stress, salty food preference	Decision tree with 19 IF-THEN rules (e.g., Rule 2: IF 34<age≤49 AND BMI≥25 AND family history of diabetes, THEN diabetic)	To screen for undiagnosed diabetes (including prediabetes) in Chinese adults using easily obtainable, non-invasive variables; to support primary care practitioners in rapid identification of high-risk individuals for targeted prevention and lifestyle intervention.

Wang Y (2019)	TG-to-HDL ratio (quartiles), ALT (IU/L, quartiles), ferritin (μg/L, quartiles), adiponectin (μg/mL, quartiles, inverse association). All biomarkers measured from plasma samples.	TG-to-HDL ratio, ALT, ferritin, adiponectin	Weighted biomarker score = (β_1×biomarker$_1$ + β_2×biomarker$_2$ + β_3×biomarker$_3$ + β_4×biomarker$_4$) × 4 / (β_1+β_2+β_3+β_4), where biomarkers are ordinal (0–3) based on quartiles.	To improve prediction of incident T2D in Chinese adults beyond traditional risk factors; to identify high-risk individuals for targeted lifestyle or pharmacological interventions.
Wu J (2019)	Age group (45–49, 50–54, 55–59, 60–70 years), family history of diabetes (first-degree relatives, yes/no), hypertension (SBP≥140 mmHg or DBP≥90 mmHg or history), overweight/obesity (BMI≥25 kg/m^2), central obesity (WC≥90 cm men, ≥85 cm women).	BMI, WC, hypertension	Points-based scoring system: age 50–54=1, 55–59=2, 60–70=3; family history=4; hypertension=3; overweight=2, obesity=4; central obesity=1.	To screen for prevalent undiagnosed diabetes in middle-aged and elderly Chinese using non-invasive, self-assessable variables

Arellano-Campos O (2019)	Office-based model: age (>40 vs. ≤40 years), family history of T2D in first-degree relatives (yes/no), waist-to-height ratio (WHtr >0.5 vs. ≤0.5), hypertension (SBP ≥140 mmHg or DBP ≥90 mmHg or treatment), BMI (≥30 vs. <30 kg/m²). Biochemical model: age (>40 years), fasting TG (>150 vs. ≤150 mg/dL), fasting plasma glucose (FPG 100–110 mg/dL, FPG 111–125 mg/dL vs. <100 mg/dL), hypertension (as above), abdominal obesity (IDF criteria: WC ≥90 cm men, ≥80 cm women).	BMI, WHtr, hypertension, TG, FPG	Points-based scoring system: points assigned based on β-coefficient categories (<0.35 =1, 0.35–0.7 =2, >0.7 =3).	To predict 3-year risk of incident T2D in middle-aged Mexican adults; office-based model for initial screening in primary care without blood tests; biochemical model for more accurate risk assessment when laboratory data are available.
Cai X (2020)	Age (years), BMI(kg/m²), FPG (mmol/L), TC (mmol/L).	BMI, FPG, TC	Nomogram: each continuous variable corresponds to points on a 0–100 scale.	To predict 5-year risk of incident T2D in Chinese adults with hypertension; to provide

				a simple, visual, and individualized risk assessment tool.
Bahijri S (2020)	Sex (male vs. female), age (<35, 35–44, 45–54, 55–64, ≥65 years), WC (WC, <94 cm men / <80 cm women [ref], 94–102 cm men / 80–88 cm women, >102 cm men / >88 cm women), history of hyperglycemia (yes/no, self-reported previous high blood glucose during medical check-up, illness, or pregnancy), family history of diabetes (none [ref], grandparent/uncle/aunt/cousin, biological parent/sibling).	WC	Points-based scoring system: points assigned based on β-coefficient categories.	To screen for undiagnosed prediabetes and T2D in the Saudi population; to provide a simple, non-invasive tool for community screening in primary care settings.
Hu H (2020)	Age (years), sex (male/female), smoking status (never, current, former), physical activity (yes/no based on ≥20 min regular	BMI, TG, FPG, physical	Cox proportional hazards regression coefficients; optimal cut-offs for	To predict 5-year risk of incident T2D in middle-aged and elderly Chinese adults

	exercise per week), BMI (kg/m², <24/≥24), TG (mmol/L, <1.1/≥1.1), FPG (mmol/L, <5.89/≥5.89).	activity, smoking	continuous predictors determined by Youden index.	who are at relatively low risk (without major underlying diseases).
Liu X (2020)	Comprehensive nomogram: age (18–29, 30–39, 40–49, 50–59, 60–69, ≥70 years), education level (Master's/higher, university, high school, low/no qualifications), parental history of diabetes (yes/no), waist-hip ratio (WHR, sex-specific quartiles), body mass index (BMI, <24 [normal], 24–27.9 [overweight], ≥28 [obese] kg/m²), sleep duration (>8 [ref], 6–8, <6 hours/night), fasting plasma glucose (FPG, <5.6 [normal], ≥5.6 mmol/L [impaired]). Concise nomogram: all above except FPG.	BMI, WHR, FPG, sleep duration	Nomogram with points assigned based on logistic regression coefficients	To predict 7-year risk of incident T2D in Chinese adults; to identify high-risk individuals for targeted lifestyle interventions, with special emphasis on sleep duration as a novel modifiable factor.

Wilkinson L (2020)	Age (years), sex (male/female), race (black/white), BMI (kg/m²), TG (mmol/L), HDL cholesterol (mmol/L), SBP (SBP, mmHg), diastolic blood pressure (DBP, mmHg), blood glucose (mmol/L).	BMI, TG, HDL, SBP, DBP, blood glucose	Bayesian logistic regression equation with coefficient estimates; formula for calculating individual 10-year T2D probability (provided in supplementary material)	To predict 10-year risk of incident T2D in black and white American adults; to identify high-risk individuals who would benefit most from weight-loss interventions; to provide a quantitative, race- and sex-specific risk assessment tool for clinical decision-making.
Zhang L (2020)	60 variables covering socio-demographics (age, sex, education, income, marital status), lifestyle (sweet flavor, high-fat diet, etc.), anthropometric (WC, BMI, WHR, etc.), laboratory (FPG, insulin, urine glucose, urine protein, creatinine, uric acid, lipids, etc.), and medical history	Sweet flavor, high-fat diet, WC, BMI, WHR, FPG, insulin,	Multiple machine learning models (LR, CART, ANN, SVM, RF, GBM); no simplified points-based score; variable importance ranking provided.	To screen for prevalent type 2 diabetes in rural Chinese adults using comprehensive clinical and lifestyle data; to identify novel risk factors (e.g., sweet flavor, urine glucose) for further investigation.

	(family history of T2DM, hypertension, CHD).	urine glucose, urine protein, creatinine, uric acid, lipids, hypertension, heart rate, pulse pressure, and others		
Shao X (2020)	Model A: age (10-year groups), sex, ethnicity, hypertension (yes/no), smoking (yes/no), alcohol intake (yes/no), WC (cm), BMI (kg/m²).	BMI, WC, smoking, alcohol, physical	Nomogram for each model; points assigned based on logistic regression coefficients	To predict 10-year risk of incident type 2 diabetes in Chinese adults; Model A for initial self-screening at home;

Model B: all Model A variables + education level, soft drink consumption (yes/no), tea consumption (yes/no), physical activity level (very light to very heavy), calorie intake (kcal/day), carbohydrate intake (g/day), fat intake (g/day), protein intake (g/day), triceps skinfold thickness (mm), sleep duration (hours/day). Model C: all Model B variables + LDL (mmol/L), HDL (mmol/L), TC (mmol/L), TG (mmol/L), insulin (μIU/L), FBG (mmol/L), HbA1c (%). Model D: LDL, HDL, TC, TG, insulin, FBG, HbA1c (all categorical).	activity, soft drink, tea, calorie intake, macronutrient intake, sleep duration, lipids, FBG, HbA1c	Model B for more comprehensive lifestyle-based risk assessment; Model C for most accurate risk estimation and individualized intervention planning; Model D for risk assessment when only laboratory data are available.

Wang H (2020)	Age (years), WC (cm), smoking (yes/no), fatty liver (diagnosed by ultrasonography, yes/no), FBG (mmol/L, <5.6 [ref], ≥5.6 to <6.1), HbA1c (%, <5.7 [ref], ≥5.7 to <6.5).	WC, smoking, fatty liver, FBG, HbA1c	Nomogram with points assigned based on Cox regression coefficients.	To predict incident type 2 diabetes in healthy Japanese adults; to identify high-risk individuals for early lifestyle interventions (diet, exercise, weight loss) and clinical monitoring.
Yang T (2020)	Age (years), sex (male/female), waistline (cm), sagittal abdominal diameter (cm), relative leg length (ratio of upper leg length to height × 100), 60-second pulse (beats/min), smoking (yes/no), alcohol (yes/no), hypertension (SBP ≥130 mmHg or DBP ≥80 mmHg, yes/no), family history of diabetes (yes/no), general health condition (self-reported, ordinal), control or loss of weight (yes/no).	Waistline, sagittal abdominal diameter, smoking, alcohol, hypertension, general health condition,	Multiple machine learning models (LDA, RF, SVM) with easy ensemble methods; interactive screening website available at http://112.126.70.33/diabetes	To screen for prevalent undiagnosed type 2 diabetes in US adults using only non-invasive, easily obtainable variables; to enable large-scale, low-cost population screening and secondary prevention.

		weight control		
Del Razo-Olvera FM (2020)	History of gestational diabetes (GDM, yes/no), BMI(kg/m²), hyperTGmia (fasting TG ≥150 mg/dL, yes/no), FPG (mg/dL).	BMI, TG, FPG, GDM (preventable through lifestyle)	Points-based scoring system: points assigned based on Cox regression coefficients using Framingham method.	To predict short-term (2.4-year) risk of incident type 2 diabetes in Mexican women of reproductive age; to address the generational impact of diabetes in pregnancy on women and their children.
Katsimpris A (2021)	Potatoes (g/d), mushrooms (g/d), onions and garlic (g/d), whipping cream (g/d), cookies (g/d), soft drinks (g/d), coffee (g/d), tomato sauce (g/d), mayonnaise (g/d), xylitol (g/d), galactose (g/d). All	All 11 dietary predictors	Logistic regression equation.	To predict prevalent type 2 diabetes using only nutrition data; to identify dietary patterns associated with T2DM for further etiological

	variables are continuous measures of usual dietary intake derived from food frequency questionnaire and 24-h food lists.			research and potential dietary interventions.
Rhee SY (2021)	Age (years), sex (male/female), BMI (kg/m²), SBP (mmHg), DBP (mmHg), FBG (mg/dL), TC (mg/dL), hemoglobin (g/dL), AST (IU/L), ALT (IU/L), GGT (IU/L), proteinuria (dipstick, yes/no), smoking (current vs. no), alcohol (drinker vs. non-drinker), exercise (yes/no), personal history (hypertension, heart disease, stroke, other [including malignancy], yes/no), family history (hypertension, heart disease, stroke, DM, other [including malignancy], yes/no).	BMI, SBP, DBP, FBG, TC, hemoglobin, AST, ALT, GGT, proteinuria, smoking, alcohol, exercise	Deep learning (RNN-LSTM) algorithm; time-dependent survival probabilities provided for each year up to 10 years; no simplified scoring system.	To predict 10-year risk of incident type 2 diabetes in Korean adults using routinely collected health check-up data; to provide personalized risk assessment integrated into the national health screening program.

Awad SF (2021)	Age (15–19, 20–24, …, 75–79 years), sex (male/female), obesity (BMI ≥30 kg/m², yes/no), smoking (current daily tobacco use, yes/no), physical inactivity (<150 min moderate or <75 min vigorous activity/week, i.e., <600 MET-minutes/week, yes/no).	Obesity, smoking, physical inactivity	Points-based scoring system: points assigned based on β-coefficients ×10.	To screen for undiagnosed prevalent type 2 diabetes in the Qatari population; to provide a simple, non-invasive tool for community screening and targeted lifestyle counselling in high-burden settings.
Cai X (2021)	Age (years), fatty liver (diagnosed by ultrasonography, yes/no), GGT (IU/L), TG (mmol/L), HbA1c (mmol/mol), FPG (mg/dL).	Fatty liver, GGT, TG, HbA1c, FPG	Nomogram with points assigned based on Cox regression coefficients.	To predict 5-year risk of incident type 2 diabetes in non-obese Japanese adults; to provide a simple, visual, and individualized risk assessment tool.

Wang Y (2021)	Age (years), family history of T2DM (yes/no), WC (cm), TC (mmol/L), TG (mmol/L), BMI (kg/m²), HDLc (mmol/L), hypertension (yes/no, defined by SBP ≥140 mmHg or DBP ≥90 mmHg or previous diagnosis).	WC, TC, TG, BMI, HDLc, hypertension	Sex-specific nomograms with points assigned based on logistic regression coefficients.	To screen for undiagnosed prevalent type 2 diabetes in adults in Western Xinjiang, China; to provide a simple, affordable, and widely implementable tool for community-based screening in underserved regions.
Cai X (2021)	Smoking status (never [ref], past, current), WC (cm), FPG (mg/dL), HbA1c (mmol/mol).	Smoking, WC, FPG, HbA1c	Nomogram with points assigned based on Cox regression coefficients.	To predict 8-year risk of incident type 2 diabetes in Japanese adults with NAFLD.
Tavares DPS (2021)	History of CVD (yes/no), history of hypertension (yes/no), sex (male/female), age (years), BMI (kg/m²), TG level (mg/dL), glucose level (mg/dL, estimated from HbA1c). The model includes interaction terms: CVD × age,	BMI, TG, glucose, hypertension	Logistic regression equation with coefficients for main effects and interaction terms.	To estimate the probability that an individual currently has type 2 diabetes based on routinely collected clinical data; to support public health policy and resource allocation

	CVD × glucose, HTA × age, HTA × BMI, glucose × HTA, sex × TG.			in remote island communities.
Mustafina SV (2021)	Full model (sex-specific): Males: FPG (≥6.0 vs. <6.0 mmol/L), BMI (≥27 vs. <27 kg/m²), HDL-C (≤0.9 vs. >0.9 mmol/L), TG (≥1.4 vs. <1.4 mmol/L), BP (≥150/90 vs. <150/90 mmHg). Females: WC (≥95 vs. <95 cm), FPG (≥5.7 vs. <5.7 mmol/L), TG (≥1.5 vs. <1.5 mmol/L), BP (≥135/90 vs. <135/90 mmHg), family history of T2DM (yes/no), BMI (≥32 vs. <32 kg/m²). Simplified model: WC (≥95 cm vs. <95 cm), history of hypertension (yes/no), family history of T2DM (yes/no).	Full model: Males: BMI, HDL-C, TG, BP, FPG; Females: WC, TG, BP, BMI, FPG. Simplified model: WC, hypertensi	Full model: Sex-specific points-based scoring system. Simplified model: Points-based scoring system.	To predict 10-year risk of incident T2DM in Russian adults aged 45-69 years; full model for primary care with laboratory data; simplified non-invasive model for self-assessment in resource-limited setting.

		on.		
Deberneh HM (2021)	FPG (mg/dL), HbA1c (%), TG (mg/dL), BMI (kg/m²), gamma-GTP (IU/L), age (years), uric acid (mg/dL), sex (male/female), smoking (currently smoking regularly/never/quit), drinking (number of days consumed alcohol), physical activity (number of days with ≥20 min exercise), family history of diabetes (parents/siblings, yes/no).	BMI, TG, gamma-GTP, uric acid, smoking, drinking, physical activity	Machine learning models (LR, RF, SVM, XGBoost, ensemble) with no simplified points-based score or regression equation provided.	To predict the occurrence of T2D in the following year (Y+1) as normal, prediabetes, or diabetes; to provide clinicians and patients with predictive information for early intervention.
Oh TJ (2021)	KDR (sex-specific): Age (40–44, 45–49, 50–54, 55–59, 60–64, 65–69 years; points vary by sex), living area (rural=0, urban=men +17, women +19), smoking (never=0, former=men +8/women +9, current=men +14/women +13), hypertension (normotensive=0,	living area, WC, hypertension, smoking	KDR: Sex-specific points-based scoring system with corresponding 10-year risk table. KDR+FPG and KDR+HbA1c: Similar scoring systems with laboratory variables.	To provide a self-assessable diabetes risk score to predict 10-year risk of incident diabetes in Korean adults; to identify high-risk individuals for targeted prevention; can be used without laboratory tests

	prehypertension=men +11/women +10, hypertension=men +20/women +19), family history of T2D (no=0, yes=men +12/women +15), WC (men <90 cm=0, ≥90 cm=+12; women <85 cm=0, ≥85 cm=+14). KDR + FPG: KDR variables + FPG KDR + HbA1c: KDR variables + HbA1c			(KDR) or with FPG/HbA1c for improved accuracy.
Stiglic G (2021)	Demographic: gender, age. Physical health: long-term illness, heart attack, hypertension, high cholesterol, stroke, chronic lung disease, asthma, arthritis, osteoporosis, cancer, stomach/duodenal ulcer, Parkinson disease, cataracts, hip fracture. Activities/quality of life: ADL (13 variables), BMI, CASP score, EURO-D depression (14 variables), cognitive	BMI, alcohol consumption, physical activity, smoking, hypertension, high	LASSO logistic regression model with coefficients; global model includes country indicators; variable selection frequency and coefficient signs provided in Supplement 1.	To predict 10-year risk of T2DM in European adults aged ≥50 years using survey-based data; to demonstrate feasibility of cross-national survey data for risk prediction; to highlight importance of model calibration across different

	abilities (2 variables). Lifestyle: smoking (ever smoked), alcohol consumption (days/week in last 6 months), vigorous physical activity, moderate physical activity.	cholesterol		populations.
Al-Lawati JA (2021)	Gender (male/female), age (≤40/≥40 years), education (ever attended school yes/no), marital status (single/married/divorced or separated/widowed), place of residence (rural/urban), combined hypertension (yes/no), BMI (<25/25–29.9/≥30 kg/m²), tobacco use (current user/ex or non-user).	place of residence, BMI, hypertension, tobacco use	Points-based scoring system (β coefficients ×5, rounded to nearest integer)	To screen for undiagnosed prevalent T2DM in the Omani population; to provide a simple, non-invasive, self-administered tool for community screening in primary care settings.
Sulistiowati E (2021)	Age group (25–39, 40–49, ≥50 years), obesity (BMI ≥25 kg/m², yes/no), central obesity (WC men ≥90 cm, women ≥80 cm,	Obesity, central obesity,	Bogor Diabetes Risk Prediction (BDRP) Chart: color-coded chart with age	To provide a simple, non-invasive, self-administered screening

	yes/no), hypertension (SBP ≥140 mmHg or DBP ≥90 mmHg or medication, yes/no), physical activity (<600 MET/week, yes/no).	hypertension, physical activity	groups on x-axis and number of risk factors (0–5) on y-axis.	tool for predicting 6-year risk of T2DM in Indonesian adults; to be used in community health centers (Posbindu) and primary care.
Tan C (2022)	WC (cm), HbA1c (%), FPG (mg/dL), HDL (mg/dL), TG (mg/dL), smoking status (never/past/current).	WC, FPG, HDL, TG, smoking	Nomogram; online calculator available (https://xy2yyjzyxk.shinyapps.io/T2DmOBESITY/).	To predict 3-year and 6-year risk of incident T2DM in Japanese adults with abdominal obesity.
Wang Y (2022)	Hypertension (yes/no), impaired fasting glucose (FPG 5.6–6.9 mmol/L, yes/no), WC (≥90 cm men/≥80 cm women, yes/no), age, sex, education, occupation, family income, smoking, drinking, physical activity, daytime napping, snoring, self-rated health, heart rate, lipid-lowering drugs, coronary heart	Hypertension, WC, BMI, smoking, drinking, physical activity, TG,	Bayesian network model with directed acyclic graph; post-test probability table.	To predict 4-year risk of new-onset diabetes in older Chinese adults using a minimal set of easily measured predictors; to visually display causal pathways and interactions among risk factors for better

	disease history, family history of diabetes, BMI, WHR, TG, HDL-C, LDL-C.	HDL-C, LDL-C		understanding of disease mechanisms.
Nicolaisen SK (2022)	Baseline HbA1c (mmol/mol), age (years), sex (male/female), BMI (kg/m²), antihypertensive drug use (yes/no), pancreatic disease (yes/no), cancer (yes/no), self-reported unhealthy diet (yes/no), doctor's advice to lose weight or change dietary habits (yes/no), lack of social support (never/almost never vs often/mostly/sometimes), good self-rated health (good vs fair/poor or excellent/very good).	BMI, antihypertensive drug use, diet, doctor's advice (behavioral), social support, self-rated health	Fine-Gray competing risk regression model with subdistribution hazard ratios; 5-year risk prediction using cumulative incidence functions.	To predict 5-year risk of progression from HbA1c-defined pre-diabetes to type 2 diabetes in Danish adults; to identify high-risk individuals for targeted preventive interventions; to inform clinical decision-making and resource allocation.

Zhao Y (2022)	Age (20–24, 25–34, 35–44, 45–54, 55–64, 65–74 years), sex (male/female), BMI (<24, 24–27.9, ≥28 kg/m²), WC (<75/75–79.9/80–84.9/≥85 cm for women; <80/80–84.9/85–89.9/≥90 cm for men), SBP (<120, 120–129, 130–139, 140–149, 150–159, 160–169, 170–179, ≥180 mmHg), family history of diabetes (no/yes).	BMI, WC, SBP	New Chinese Diabetes Risk Score (NCDRS): points-based scoring system with each predictor category assigned points (β coefficient ×10, rounded to integer).	To predict 6-year risk of incident T2DM in Chinese adults using simple non-invasive measures; recommended as unified screening tool by Chinese diabetes guidelines.
Sadek K (2022)	Age (18–35 ref, 36–54=12, ≥55=26), gender (female ref, male=9), WHR (normal ref, abnormal ≥1 male/≥0.86 female=7), history of hypertension (yes=7), history of hyperlipidemia (yes=12), education (university ref, secondary/absent=8).	WHR, BMI, hypertension, hyperlipidemia	Points-based scoring system (β coefficients ×10).	To screen for undiagnosed T2DM and impaired glucose metabolism in the Qatari population; to provide simple, cost-effective tools for primary care and community-based screening.

Edlitz Y (2022)	Anthropometric model: age (5-year bands), sex, weight (kg), height (cm), WC (cm), hip circumference (cm), BMI (kg/m²), WHR. Four blood test model: age, sex, HbA1c (%), GGT (U/L), HDL (mmol/L), TG (mmol/L).	Anthropometric: weight, height, WC, hip circumference, BMI, WHR; Blood tests: HbA1c, GGT, HDL, TG	Anthropometric scorecard with three risk groups (1%, 5%, 9% 7-year risk); four blood test scorecard with four risk groups (0.5%, 3%, 10%, 23% 7-year risk); logistic regression and Cox regression coefficients also provided.	To predict 7-year risk of T2D onset in adults aged 40–69 using either simple anthropometric measures (accessible to general population) or four common blood tests (more accurate for clinical use).
Liu Q (2022)	Age (years), gender (male/female), education (elementary and below/junior high/technical secondary or high school/junior college and above), marital status	Exercise, smoking, drinking, WC, SBP, FPG, TC,	Machine learning models (LR, DT, RF, XGBoost) with hyperparameter tuning via Bayesian optimization; XGBoost performed best	To predict 2-year risk of incident T2DM in Chinese elderly adults (≥65 years) using health screening data.

	(married/divorced/widowed/single), hypertension (yes/no), exercise (yes/no), current smoking (yes/no), current drinking (yes/no), WC (cm), SBP (mmHg), FPG (mmol/L), TC (mmol/L), TG (mmol/L), HDL-C (mmol/L), LDL-C (mmol/L), ALT (U/L), AST (U/L), TBIL (μmol/L), Scr (μmol/L), BUN (mmol/L), uric acid (μmol/L).	TG, HDL-C, LDL-C, ALT, AST, TBIL, Scr, BUN, uric acid	(AUC 0.7805); SHAP for feature importance visualization.	
Martínez-Hervás S (2022)	FPG (mg/dL, ≤92, 92–106, >106); fasting TG (mg/dL, ≤97, 97–180, >180); age (years, >56); BMI (kg/m², >28.24; gender).	FPG, fasting TG, BMI	CHAID decision tree diagram with hierarchical nodes: Node 1 (FPG ≤92) further split by FTG and age; Node 2 (FPG 92–106) split by age and BMI; Node 3 (FPG >106) terminal high-risk node (≥25%	To predict 7.5-year risk of incident T2DM in general population using simple clinical and laboratory data; to identify high-risk individuals for targeted prevention.

			incident T2D).	
Dong W (2022)	LR model: age (years, with quadratic term), BMI (kg/m²), WHR (per 0.1), smoking status (current smoker yes/no), sleep duration (hours/day), vigorous recreational activity time (min/week), fruit consumption (servings/week), age×sleep duration interaction.	BMI, WHR, smoking, sleep duration, vigorous recreational activity time, fruit consumption	LR model equation provided; online risk calculator available (GitHub).	To screen for undiagnosed DM and pre-DM in Chinese adults using non-laboratory risk factors; to identify high-risk individuals for confirmatory blood testing in primary care; to incorporate novel modifiable factors (sleep, vigorous activity) for more accurate risk assessment.

	XGBoost model: age (years), BMI (kg/m²), WHR, SBP (mmHg), WC (cm), sleep duration (hours/day), smoking status (current smoker yes/no), vigorous recreational activity time (min/week).	BMI, WHR, WC, SBP, smoking, sleep duration, vigorous recreational activity time	XGBoost model with SHAP feature importance and partial dependence plots; online risk calculator available (GitHub).	
Tong C (2022)	Sex (male/female), age (<30/30–39/40–49/50–59/≥60 years), BMI (underweight/normal/overweight/obesity), FPG (<5.6/5.6–6.1/6.2–7.0 mmol/L), hypertension (yes/no), dyslipidaemia (yes/no), sCr (quartiles: <57.2/57.2–68.1/68.2–78.9/>78.9	BMI, FPG, hypertension, dyslipidaemia, sCr	Dynamic nomogram (online application: https://nomogramxhy.shinyapps.io/DynNomapp/) predicting 3-, 5-, and 10-year T2DM probability; risk score based on weighted Cox regression coefficients.	To predict T2DM risk in urban Chinese adults undergoing routine health check-ups; dynamic nomogram for clinicians, risk score for self-assessment by examinees.

	µmol/L).			
Boersma HE (2022)	Model 1: Age group (18-44, 45-55, 56-64, ≥65 yrs), BMI group (<25.0, 25.0-29.9, ≥30.0 kg/m²), SAF group (<1.90, 1.90-2.49, 2.50-2.89, ≥2.90 AU). Model 2: All variables from Model 1 + number of parents with diabetes (0, 1, 2).	BMI, SAF (skin autofluorescence, a marker of advanced glycation endproducts which are influenced by diet and smoking)	Points-based scoring system (regression coefficients provided for each category).	To provide a fast, non-invasive screening tool for identifying individuals at high risk for undiagnosed and incident T2DM, suitable for medical and non-medical settings.

Wang S (2023)	Age, drinking, consumption frequency of cereals, potatoes, beans, fruits, eggs, dairy, poultry and fish, DBP, FPG, TC, TG, HDL-C, LDL-C.	Drinking, diet (cereals, potatoes, beans, fruits, eggs, milk, poultry, fish), DBP, FPG, TC, TG, HDL-C, LDL-C.	Logistic regression equation; Machine learning models (BPNN, SVM, CART, C4.5, DNN) with performance metrics.	To provide a methodological reference for predicting T2DM risk; to compare the performance of six machine learning models and identify the most effective one for clinical application.
Zheng M (2023)	Sex-specific models: Age (years), BMI (kg/m²), WC (categorical), DBP (categorical), TG (≥2.3 mmol/L), FPG (mmol/L), fatty liver	BMI, WC, DBP, TG, FPG, fatty liver	Points-based scoring system with separate score charts for men and women; Web calculator available at	To predict 5-year risk of new-onset T2D in Chinese adults with IFG; to identify high-risk individuals for

	(absent/mild/moderate/severe).		https://diabetesprediction.shinyapps.io/RP-IFG/.	targeted lifestyle interventions and clinical monitoring.
Yonel Z (2023)	Age (years), sex (male/female), BMI (kg/m²), smoking status (never/former/current), first-degree relative with diabetes (yes/no), presence of dental prosthesis (yes/no), presence of self-reported mobile teeth (yes/no), history of periodontal treatment in last 5 years (yes/no), number of sites with probing pocket depth ≥5 mm (0 sites/up to 2 sites/3 or more sites).	BMI, smoking status, mobile teeth, periodontal treatment, probing pocket depths	Logistic regression equation with coefficients; Simplified point-based scoring system with score chart.	To identify individuals with undiagnosed NDH or T2D in dental settings; to be used as a first-stage risk assessment tool, with those above threshold offered point-of-care HbA1c testing.

Hu Y (2023)	Age (years), BMI (kg/m²), FPG (mmol/L), TG (mmol/L), SBP (mmHg), Scr (μmol/L).	BMI, FPG, TG, SBP, Scr	Nomogram; Risk prediction formula: Predicted risk (5-yr) = 1 − S_0 (5-yr)^exp(LP), where LP = 1.5168×FPG + 0.0182×age + 0.0482×BMI + 0.0096×SBP + 0.0665×TG + 0.0062×Scr, S_0 (5-yr) = 0.999993.	To provide a personalized, simple tool for clinicians to identify individuals at high risk of developing prediabetes within 5 years, enabling early lifestyle interventions.
Yang J (2024)	Age (40-54/55-74/≥75 years), gender (male/female), BMI (24-28/≥28 kg/m²), FPG (mmol/L, continuous), HDL-C (per 0.1 mmol/L), LDL-C (per 0.1 mmol/L), fatty liver (yes/no), liver dysfunction (ALT>50 & AST>40, yes/no).	BMI, FPG, HDL-C, LDL-C, fatty liver, liver dysfunction	Nomogram; Logistic regression equation: ln(p/1-p) = 0.181×age(40-54)/0.973×age(55-74)/1.868×age(≥75) − 0.192×gender(male) + 0.151×FPG − 0.538×BMI(24-28) − 0.538×BMI(≥28) −	To predict 3-year risk of progression from prediabetes to diabetes; to identify high-risk patients for targeted lifestyle interventions and clinical management.

			0.109×HDL-C + 0.021×LDL-C + 0.365×fatty liver + 0.444×liver dysfunction – 10.038.	
Ojurong be TA (2024)	WHR (continuous), TG (mmol/L), catalase (U/L), atherogenic indices of plasma (log(TG/HDL-C)).	WHR, TG, catalase, atherogenic indices of plasma	Web-based application (https://iv3p9h-nurudeen-adegoke.shinyapps.io/Diabetic/); Logistic regression coefficients.	To enable early detection of type 2 diabetes in resource-limited settings.
Ghiasi Hafezi S (2024)	Logistic regression: BMI (categorical), VAI (continuous), LAP (categorical), AVI (continuous). Decision tree: BMI, LAP, VAI (with hierarchical splitting rules).	BMI, VAI, LAP, AVI	Logistic regression coefficients; Decision tree with CHAID algorithm (14 IF-THEN rules).	To predict 10-year risk of T2DM using simple, affordable anthropometric indices suitable for large-scale screening without advanced technology or specialized

				skills.
Xu J (2024)	Model 1: Sex (male/female), BMI (kg/m²), family history of DM (yes/no), DBP (mmHg). Model 2: Family history of DM, HbA1c (%). Model 3: Family history of DM, HbA1c (%), FPG (mg/dL).	BMI, DBP, HbA1c, FPG	Logistic regression equations (coefficients provided in Table 3); Excel spreadsheet for risk calculation.	To predict 5-year risk of incident T2DM in Japanese adults; non-invasive model suitable for large-scale screening; invasive models for more accurate risk assessment.
Lin Y (2024)	Age (years), gender (male/female), BMI (kg/m²), FBG (mmol/L), TG (mmol/L), ALT (U/L), ALT/AST ratio, BUN (mmol/L), Hb (g/L).	BMI, FBG, TG, ALT, ALT/AST, BUN, Hb	Dynamic online nomogram (https://diabetes-prediction.shinyapps.io/DiabetesDynNom/); Logistic regression coefficients.	To identify older adults with normal fasting glucose who are at risk of diabetes, enabling early detection in primary healthcare settings without requiring OGTT or HbA1c tests.

Xie R (2025)	Clinical CDRS: Age (years), sex (male/female), BMI (kg/m²), smoking status (never/ex/current), family history of diabetes (none/parent or sibling/parent and sibling), antihypertensive medication (yes/no), steroid medication (yes/no), HbA1c (%). Additional 11 metabolites: glucose, citrate, lactate, pyruvate, 3-hydroxybutyrate, acetate, glutamine, tyrosine, IDL-CE-pct, M-LDL-TG-pct, LA-pct.	BMI, smoking, antihypertensive medication, HbA1c, glucose, and metabolites reflecting diet/lipid metabolism	Cox regression coefficients; risk score equation can be constructed.	To predict 10-year risk of incident type 2 diabetes; to identify high-risk individualsusing high-throughput nuclear magnetic resonance metabolomics in clinical settings.
Kawasoe S (2025)	Model 1 (invasive): Age (years, continuous), sex (male=1, female=0), BMI (kg/m², continuous), FBG (mg/dL, continuous), HbA1c (%, continuous). Model 2 (non-invasive): Age (years,	BMI, FBG, HbA1c, hypertension	Logistic regression equations: Model 1: p = 1/(1+exp(-(-30.3726 – 0.0184×age + 0.2876×sex + 0.0905×BMI + 0.1101×FBG +	To predict 5-year risk of incident diabetes using routine health check-up data.

	continuous), sex (male=1, female=0), BMI (kg/m², continuous), hypertension (yes=1, no=0).		2.5471×HbA1c))); Model 2: p = 1/(1+exp(-(8.4968 + 0.0200×age + 0.5096×sex + 0.1522×BMI + 0.4644×hypertension)))	
Wang Y (2025)	Age (years), WC (cm), ALT (IU/L), HbA1c (%), FPG (mg/dL), fatty liver (yes/no), current smoking (yes/no), heavy alcohol consumption (yes/no), HDL-C (mg/dL, protective).	WC, ALT, HbA1c, FPG, fatty liver, smoking, alcohol consumption, HDL-C	Cox regression HRs; Combined risk model with ROC-derived cutoff (predicted probability >0.017) for high risk.	To predict 8-year risk of incident T2DM; to identify high-risk individuals for targeted lifestyle interventions.
Shahraki PK (2025)	Diabetes Model 1: WHR (≥0.795=1 point), HbA1c (≥4.8 mg/dL=1 point), OGTT-AUC (≥773=1 point), lipid index (≥0.045=1 point).	WHR, HbA1c, OGTT-AUC, FPG,	Points-based scoring system; risk scores calculated as sum of individual scores for each significant risk factor.	To predict 13-year risk of developing diabetes and prediabetes in first-degree relatives of T2DM patients.

	Diabetes Model 2: WHR, HbA1c, FPG (≥88.5 mg/dL=1 point), lipid index.	SBP, WC, lipid index (composite of TG, cholesterol, HDL, LDL)		
Chee HK (2025)	Model IV: BMI ≥23 kg/m² (yes/no), family history of DM (yes/no), smoking (yes/no), Stage III/IV periodontitis (yes/no). Model V: BMI ≥23 kg/m², family history of DM, smoking, severe periodontitis (yes/no).	BMI, smoking, periodontitis	Logistic regression coefficients; risk model based on ORs.	To help dentists identify patients at high risk for undiagnosed DM and prediabetes using simple clinical indicators.

Munekawa C (2025)	Age (years), sex (male=1, female=0), BMI (kg/m²), SBP (mmHg), log-TG (log of TG, mg/dL), HDL (mg/dL), log-ALT (log of alanine aminotransferase, U/L), FPG (mg/dL), weight gain (yes/no, defined as weight gain ≥10 kg vs. age 20), smoking status (yes/no).	BMI, SBP, TG, HDL, ALT, FPG, weight gain, smoking	Cox regression equation: 10-year risk = 1 – (1 – 0.00000000075)^exp(βx), where βx = 0.032×age – 0.335×male + 0.071×BMI + 0.003×SBP + 0.076×log-TG – 0.009×HDL + 0.437×log-ALT + 0.140×FPG + 0.071×weight_gain + 0.466×smoking.	To predict 10-year risk of incident T2DM in Japanese adults using routinely collected health check-up data.
Kivinie mi E (2025)	Baseline model: age (years), sex, BMI (kg/m²), WC (cm), SBP (mmHg), diastolic blood pressure (mmHg), HDL-C (mmol/L), TC (mmol/L), TG (mmol/L), glucose (mmol/L), current smoking (yes/no), BP medication (yes/no), lipid medication (yes/no), diabetes family	BMI, WC, SBP, HDL-C, TC, TG, glucose, smoking, BP	Logistic regression equations.	To predict 5- and 8-year risk of incident T2D; to assess the incremental predictive value of circulating metabolic biomarkers beyond established clinical risk factors.

	history (yes/no). Extended models: baseline variables + 154 NMR-based metabolites (e.g., glycine, docosahexaenoic acid, glutamine, histidine, tyrosine).	medication, lipid medication, and metabolites reflecting diet and metabolic health		
Alareeki A (2025)	Age group (10-year intervals), sex (male/female), obesity (BMI ≥30 kg/m²), smoking (current daily tobacco use), physical inactivity (<150 min moderate or <75 min vigorous activity/week)	Obesity, smoking, physical inactivity	Points-based scoring system (β coefficients ×10, rounded to nearest integer); scoring table and equation provided	To identify individuals at high risk of undiagnosed T2DM in Jordan
Chen YE (2025)	HbA1c (%), blood glucose (mg/dL), peak expiratory flow (PEF, L/sec), PRS (standardized), age (years), WHR, family	HbA1c, blood glucose,	Logistic regression coefficients; risk prediction equation.	To predict 3-year risk of incident T2DM in Han Chinese; applicable to East

	history of diabetes (yes/no), BMI (kg/m²), TG (mg/dL), serum glutamate pyruvate transaminase (SGPT, U/L), prediabetes (ICD code), self-reported hyperlipidemia (yes/no), mixed hyperlipidemia (ICD code, yes/no).	PEF, WHR, BMI, TG, SGPT, prediabetes, hyperlipidemia		Asian populations with cross-population validation.
Han X (2025)	Gender (male:1 point), age (60-69:1, 70-79:3, ≥80:5 points), smoking (yes:3 points), central obesity (yes:5 points), exercise (no:2 points), BMI (24.0-27.9:5, ≥28:9 points), WHtR (>0.5:9 points), hypertension (yes:3 points), fatty liver (yes:9 points), family history of diabetes (yes:5 points), FBG (100-110:12, >110:24 points), low HDL-C (yes:1 point), TG (≥1.7:6 points), high uric acid (yes:2	Smoking, exercise, BMI, WHtR, hypertension, fatty liver, FBG, TG, HDL-C, uric acid,	Points-based scoring system (iCARE-DM); total score calculated as sum of points for each risk factor.	To predict 5-year risk of incident T2DM in elderly Chinese adults using routine public health service data.

	points), TYG (8.25-8.6:5, 8.6-9.0:9, ≥9.0:14 points).	TYG		
Zhang X (2025)	Age (years), exercise (yes/no), WC (cm), FPG (mmol/L), TG (mmol/L), SBP (mmHg), dessert intake (high/low).	Exercise, WC, TG, SBP, dessert intake	Nomogram; Cox regression equation: Log(Hazard) = –0.7077×exercise(yes) + 0.0104×SBP – 0.0345×age + 0.0186×waist + 0.1933×TG + 0.5859×FPG + 0.5622×dessert intake(high).	To predict 3-year risk of progression from prediabetes to diabetes in older adults.
AlKetbi LM (2025)	Age (years), random blood glucose (mmol/L), BMI (kg/m²), current smoking (yes/no), sex (male/female), HBA1C (%), mean BP (mmHg), eGFR percentile (3rd, 10th, 50th, 75th, 90th, 95th, 97th), HDL	Random blood glucose, BMI, smoking,	Cox regression coefficients (HRs with 95% CI); risk stratification using identified predictors.	To predict 12-year risk of incident T2DM in Emirati adults; to inform clinical practice and guide targeted preventive interventions.

	(mmol/L), hypertension at baseline (yes/no).	mean BP, HDL, hypertension		
Satoh T (2025)	Age (40-<50, 50-<60, 60-<70, 70-<80, ≥80 years), sex (male), BMI (18.5-<22, 22-<25, 25-<30, ≥30 kg/m²), SBP (120-<130, 130-<140, 140-<150, 150-<160 mmHg), TG (100-<150, ≥150 mg/dL), HDL-C (<40 mg/dL), LDL-C (120-<140, ≥140 mg/dL), AST (30-<50, ≥50 IU/L), γ-GTP (50-<100, ≥100 IU/L), HbA1c (5.5-<6.0, 6.0-<6.5%), eGFR (45-<60 G3a, 30-<45 G3b mL/min/1.73 m²), urine protein (±, ≥+), antihypertensive medication (yes), dyslipidemia medication (yes), heavy drinking (yes), exercise habits (yes).	BMI, SBP, TG, HDL-C, LDL-C, AST, γ-GTP, HbA1c, eGFR, urine protein, antihypertensive medication,	Points-based scoring system (coefficients multiplied and rounded to integer weights); cumulative incidence curves by score category with 1-, 3-, and 5-year risk estimates.	To predict risk of incident T2DM using routine health check-up data; to enable risk stratification and guide preventive interventions in clinical and public health practice.

	dyslipidemia medication, heavy drinking, exercise habits

ALT, alanine aminotransferase; ANN, artificial neural network; AST, aspartate aminotransferase; AUC, area under the curve; AVI, abdominal volume index; BMI, body mass index; BP, blood pressure; BPNN, back propagation neural network; BUN, blood urea nitrogen; CART, classification and regression tree; CASP, Control, Autonomy, Self-Realization and Pleasure; CHAID, Chi-squared Automatic Interaction Detector; CI, confidence interval; CV, cross-validation; CVD, cardiovascular disease; DBP, diastolic blood pressure; DM, diabetes mellitus; DNN, deep neural network; eGFR, estimated glomerular filtration rate; EURO-D, EURO-D depression scale; FPG, fasting plasma glucose; GBM, gradient boosting machine; GDM, gestational diabetes mellitus; GGT, gamma-glutamyl transferase; GRS, genetic risk score; Hb, hemoglobin; HbA1c, glycated hemoglobin; HDL, high-density lipoprotein; HDL-C, high-density lipoprotein cholesterol; HR, hazard ratio; IDF, International Diabetes Federation; IFG, impaired fasting glucose; LAP, lipid accumulation product; LASSO, Least Absolute Shrinkage and Selection Operator; LDA, linear discriminant analysis; LDL, low-density lipoprotein; LDL-C, low-density lipoprotein cholesterol; LR, logistic regression; LRP,

layer-wise relevance propagation; MET, metabolic equivalent; MFP, multivariable fractional polynomial; NAFLD, non-alcoholic fatty liver disease; NR, not reported; OGTT, oral glucose tolerance test; OR, odds ratio; PCOS, polycystic ovary syndrome; PEF, peak expiratory flow; PRS, polygenic risk score; RBS, random blood sugar; RF, random forest; RNN-LSTM, recurrent neural network long short-term memory; ROC, receiver operating characteristic; SBP, systolic blood pressure; Scr, serum creatinine; SGPT, serum glutamate pyruvate transaminase; SHAP, SHapley Additive exPlanations; SMOTE, synthetic minority over-sampling technique; SNP, single nucleotide polymorphism; SVM, support vector machine; T2D, type 2 diabetes; T2DM, type 2 diabetes mellitus; TBIL, total bilirubin; TC, total cholesterol; TG, triglycerides; TYG, triglyceride-glucose index; VAI, visceral adiposity index; WC, waist circumference; WHR, waist-to-hip ratio; WHtR, waist-to-height ratio; XGBoost, eXtreme Gradient Boosting; 2hPG, 2-hour plasma glucose