JKMS

Original Article
Public Health & Preventive Medicine

Check for updates

Validation of the Framingham Diabetes Risk Model Using Community-Based KoGES Data

Hye Ah Lee [1] Hyesook Park [2,3] and Young Sun Hong [4]

[1]Clinical Trial Center, Ewha Womans University Mokdong Hospital, Seoul, Korea
[2]Department of Preventive Medicine, College of Medicine, Ewha Womans University, Seoul, Korea
[3]Graduate Program in System Health Science and Engineering, Ewha Womans University, Seoul, Korea
[4]Department of Internal Medicine, College of Medicine, Ewha Womans University, Seoul, Korea

OPEN ACCESS

ORCID iDs
Hye Ah Lee
https://orcid.org/0000-0002-4051-0350
Hyesook Park
https://orcid.org/0000-0002-9359-6522
Young Sun Hong
https://orcid.org/0000-0001-8552-9066

Disclosure
The authors have no potential conflicts of interest to disclose.

ABSTRACT

Background: An 8-year prediction of the Framingham Diabetes Risk Model (FDRM) was proposed, but the predictor has a gap with current clinical standards. Therefore, we evaluated the validity of the original FDRM in Korean population data, developed a modified FDRM by redefining the predictors based on current knowledge, and evaluated the internal and external validity.

Methods: Using data from a community-based cohort in Korea (n = 5,409), we calculated the probability of diabetes through FDRM, and developed a modified FDRM based on modified definitions of hypertension (HTN) and diabetes. We also added clinical features related to diabetes to the predictive model. Model performance was evaluated and compared by area under the curve (AUC).

Results: During the 8-year follow-up, the cumulative incidence of diabetes was 8.5%. The modified FDRM consisted of age, obesity, HTN, hypo-high-density lipoprotein cholesterol, elevated triglyceride, fasting glucose, and hemoglobin A1c. The expanded clinical model added γ-glutamyl transpeptidase to the modified FDRM. The FDRM showed an estimated AUC of 0.71, and the model's performance improved to an AUC of 0.82 after applying the redefined predictor. Adding clinical features (AUC = 0.83) to the modified FDRM further improved in discrimination, but this was not maintained in the validation data set. External validation was evaluated on population-based cohort data and both modified models performed well, with AUC above 0.82.

Conclusion: The performance of FDRM in the Korean population was found to be acceptable for predicting diabetes, but it was improved when corrected with redefined predictors. The validity of the modified model needs to be further evaluated.

Keywords: Diabetes; Framingham Offspring Study; Probability; Risk Assessment

INTRODUCTION

In 2003, the World Health Organization and International Diabetes Federation predicted that the prevalence of diabetes would reach 350 million by 2030.[1] However, in 2021, the International Diabetes Federation estimated that 537 million people already had diabetes worldwide, a number 1.5-fold higher than the official prediction, and that an additional 240

million people had undiagnosed diabetes.[2] The prevalence and disease burden of diabetes are expected to increase continuously.

A simulation study demonstrated that early detection of diabetes through screening improves the prognosis of patients.[3] Identification of individuals at high risk of disease allows effective disease prevention through targeted interventions and delays the development of disease. Therefore, several studies have evaluated predictive models for diabetes, including some based on machine learning. Fregoso-Aparicio et al.[4] systematically reviewed 90 studies of prediction models based on machine learning and deep leaning that were published between 2017 and 2021. Although several of the models demonstrated good performance, they have limited clinical applications. In 2007, the Framingham Diabetes Risk Model (FDRM), which is based on age, sex, body mass index (BMI), blood pressure, parental history of diabetes, fasting blood glucose (FBG), high-density lipoprotein cholesterol (HDL-C), and triglyceride (TG) levels, was proposed in the Framingham offspring study.[5] The FDRM is based on easily available features that have good discriminative ability (area under the curve [AUC] = 0.85). It was developed for use in middle-aged adults and predicts the 8-year diabetes risk. Since 2011, the World Health Organization has recommended the use of hemoglobin A1c (HbA1c) for the diagnosis of diabetes.[6] Additionally, in 2017, the American College of Cardiology/American Heart Association (ACC/AHA) changed its definition of hypertension (HTN) from blood pressure ≥ 140/90 mmHg to ≥ 130/80 mmHg.[7] Meanwhile, the Korean Society of HTN announced that it would maintain the existing diagnostic criteria in accordance with the 2022 HTN management guidelines.[8] These discrepancies may affect the predictive power of disease prediction models.

Therefore, we evaluated the validity of the FDRM in an independent cohort in Korea and herein propose a modified version thereof based on predictors redefined using current clinical criteria. Additionally, we assessed whether predictive performance improved when additional predictors were added to the modified FDRM.

METHODS

Data and study subjects

We obtained data from the community-based cohort of the Korean Genome and Epidemiology Study (KoGES), which was established in 2001–2002 and has a biennial follow-up. The study included 10,030 volunteers aged 40–69 years who were living in the rural area of Ansung (n = 5,018) and the industrial area of Ansan (n = 5,012) in the province of Gyeonggi-do. Detailed information on the cohort has been published elsewhere.[9]

We excluded participants with cancer, a history of cardiovascular disease (myocardial infarction, stroke, coronary artery disease, or congestive heart failure), diabetes, an FBG level ≥ 126 mg/dL, or HbA1c ≥ 6.5%. This study included data up to the fourth follow-up, conducted in 2009–2010. Data were randomly split into derived and validation datasets at a ratio of 7:3. Finally, data from 2,547 males and 2,862 females (derivation dataset) were used to validate and improve the FDRM, while data from 1,097 males and 1,221 females (validation dataset) were used for internal validation of the prediction models. For temporal validation, data from the fourth follow-up of the same cohort conducted in 2009–2010 were used as the baseline, whereas incident diabetes recorded in the fifth to eighth follow-ups (2017–2018) was used as the outcome. Based on the eligibility criteria, data from 1,914 males and 2,177 females (temporal validation data) were included in the study.

For external validation, the validity of the predictive models was assessed using data from the KoGES Cardiovascular disease association study (KoGES-CAVAS), which was selected based on the availability of data that could be explained by the developed predictive model. We evaluated the validity of the prediction model among subjects aged 40–69 years using diagnosed and undiagnosed diabetes as the outcome. Thus, after applying the eligibility criteria, we obtained data from 790 males and 1,253 females (external validation data). Briefly, KoGES-CAVAS were collected from 28,337 people living in rural areas from 2005 to 2011 for the purpose of preparing measures for disease prevention and early diagnosis through identification of risk factors related to the development of diseases, focusing on cardiovascular disease.[9] Up to the 4th follow-up (2014-2016) data are available. HbA1c at baseline was measured only in some regional participants, so only 5,906 data exist. **Fig. 1** presents an overview of the study procedures.
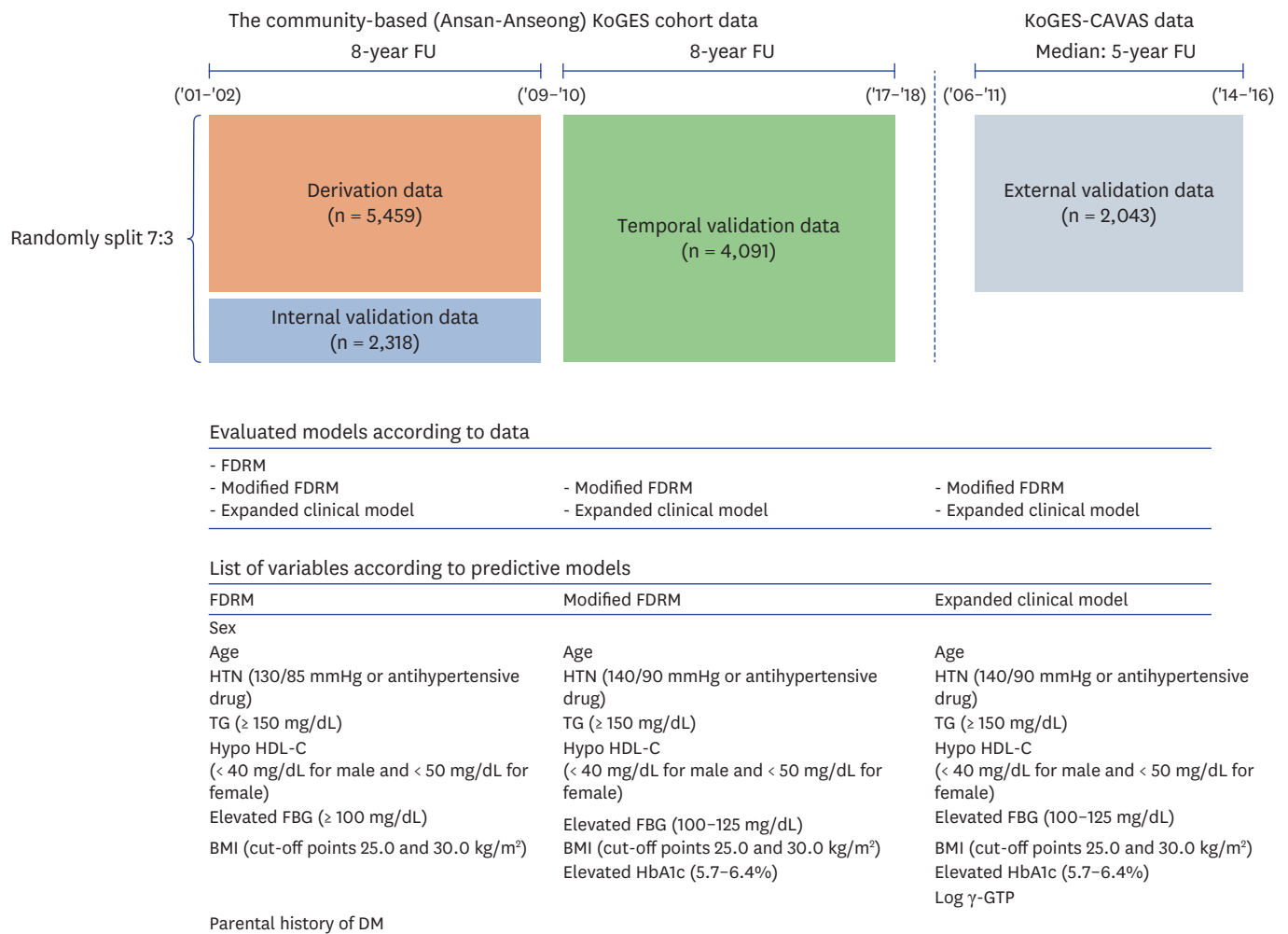
Fig. 1. Study summary.
The data sources used in this study, predictive models evaluated, and variables included therein are presented.
KoGES = Korean Genome and Epidemiology Study, KoGES-CAVAS = Korean Genome and Epidemiology Study-cardiovascular disease association study, FDRM = Framingham Diabetes Risk Model, HTN = hypertension, TG = triglyceride, HDL-C = high-density lipoprotein cholesterol, FBG = fasting blood glucose, BMI = body mass index, HbA1c = hemoglobin A1c, DM = diabetes mellitus, γ-GTP = γ-glutamyl transpeptidase.

Outcome

Incident diabetes was defined as diagnosis of diabetes by a physician and the use of medication for diabetes, or as an FBG level ≥ 126 mg/dL or HbA1c level ≥ 6.5%. The cumulative incidence of diabetes was 8.5% (255 of 2,547 males and 205 of 2,862 females) in the 8-year follow-up derivation dataset and 7.8% (85 of 1,097 males and 95 of 1,221 females) in the internal validation dataset. In the temporal validity dataset, 368 individuals (9.0%) developed diabetes during the 8-year follow-up period (172 of 1,914 males and 196 of 2,177 females). In the external validation dataset from the KoGES-CAVAS, the estimated cumulative incidence of diabetes was 6.8% (62 of 790 males and 76 of 1,253 females).

Assessment of prediction model

We assessed the validity of the FDRM and modified it as follows: first, the predicted individual diabetes risk was calculated by applying the beta coefficient derived from the FDRM to the derivation and validation datasets. Thus, according to the FDRM definition,[5] eight component features were defined: age, sex, blood pressure, parental history of diabetes, BMI, and FBG, HDL-C, and TG levels:

Beta = −5.517 + (male × −0.01) + [HTN (> 130/85 mmHg or receiving therapy) × 0.498] + [high TG (≥ 150 mg/dL) × 0.575] + [hypo HDL-C (< 40 mg/dL for males and < 50 mg/dL for females) × 0.944] + [high FBG (≥ 100 mg/dL) × 1.98] + (parental history of diabetes × 0.565)

if BMI 25.0–29.9 kg/m^2, + 0.301; if BMI ≥ 30.0 kg/m^2, + 0.92 and if age 50-64 years, −0.018; if age ≥ 65 years, −0.081

Second, we modified the FDRM by adding the revised definition of HTN (blood pressure ≥ 140/90 mmHg or use of antihypertensive drugs[8]) and HbA1c (5.7–6.4%),[6,10] as well as by excluding sex and parental diabetes history (named modified FDRM).[5,11] Third, additional clinical features were applied to the modified FDRM to improve its performance (named expanded clinical model). We assessed diabetes-related features identified in previous studies[10,12-15] that are recorded routinely and were available in the KoGES dataset. We assessed the associations of incident diabetes with 35 features (**Supplementary Table 1**). Regarding redundant features, we excluded highly correlated features (correlation coefficient [γ] > 0.8). The stepwise selection method was used to identify the most important features in the modified FDRM. To evaluate collinearity among the features included in the predictive model, we estimated the variance inflation factor. We also evaluated the interaction terms for the features included in the predictive model. The internal and external validity of the modified FDRM and expanded clinical model were evaluated.

Statistical analysis

Statistical analysis was performed using R (version 3.6.2; R Foundation for Statistical Computing, Vienna, Austria) and SAS (version 9.4; SAS Institute Inc., Cary, NC, USA) software. Demographic characteristics are presented as means with standard deviations or medians with interquartile ranges for continuous variables, as appropriate, and as numbers with percentages for categorical variables. Associations of incident diabetes with the features are presented as odds ratios, 95% confidence intervals (CIs), c-index (i.e., AUC) values, and scores derived from the logistic model. We calculated the AUC to compare the predictive models in terms of discrimination ability. Within-study validity was assessed using 10-fold cross validation. Sensitivity analysis was performed as follows. In the validation data, HTN was defined according to the ACC/AHA standards[7] (blood pressure ≥ 130/80 mmHg or use

of antihypertensive drugs) and the difference in discrimination power was evaluated when applied to the prediction models. We also evaluated model performance for diabetes, defined as physician diagnosis and diabetes medication use, FBG level ≥ 126 mg/dL or HbA1c level ≥ 6.5% or 2-hour blood glucose level ≥ 200 mg/dL in oral glucose tolerance test. Due to data availability, it was evaluated on internal and temporal validation data. We also evaluated model performance in subgroups (male/female, normoglycemia/prediabetes). A two-tailed *P* value < 0.05 was considered to indicate statistical significance.

Ethics statement

This study was performed in accordance with the Declaration of Helsinki. The study protocol was approved by the Institutional Review Board (IRB) of Ewha Womans University Hospital (approval no. EUMC 2021-03-008). The requirement for written informed consent was waived by the IRB because the study used an anonymous dataset.

RESULTS

Table 1 presents the baseline characteristics of the participants. In the derivation dataset, 47.1% of the participants were males; their mean age was 51.5 (range, 40–69) years, 40.4% had a BMI > 25.0 kg/m^2 and 7.8% had a parental history of diabetes. Approximately a quarter of the participants had a blood pressure ≥ 140/90 mmHg (24.4%) and half of the participants

Table 1. Basic characteristics of the participants in the derivation and validation datasets

Characteristics	Derivation data (n = 5,409)	Internal validation data (n = 2,318)	*P* value
Male	2,547 (47.09)	1,097 (47.33)	0.848
Age, yr	51.52 ± 8.73	51.72 ± 8.79	0.364
Parental history of diabetes	422 (7.80)	170 (7.33)	0.479
BMI, kg/m^2	24.41 ± 3.10	24.42 ± 3.06	0.916
< 25.0	3,221 (59.56)	1,372 (59.24)	0.608
25.0–29.9	1,939 (35.85)	848 (36.61)	
≥ 30.0	248 (4.59)	96 (4.15)	
Waist circumference, cm	81.99 ± 8.73	82.12 ± 8.67	0.558
≥ 94 for males and ≥ 80 for females	1,687 (31.22)	748 (32.28)	0.358
SBP, mmHg	120.54 ± 17.82	120.10 ± 18.31	0.319
DBP, mmHg	80.00 ± 11.45	79.61 ± 11.32	0.172
≥ 140/90 mmHg	1,321 (24.42)	550 (23.73)	0.514
≥ 130/80 mmHg	2,825 (52.23)	1,178 (50.82)	0.256
Antihypertensive drug use	441 (8.15)	194 (8.37)	0.751
HbA1c, %	5.55 ± 0.35	5.54 ± 0.34	0.133
5.7–6.4	2,050 (37.90)	854 (36.84)	0.379
Fasting blood glucose, mg/dL	83.06 ± 8.93	82.69 ± 8.81	0.089
100–125	258 (4.80)	101 (4.37)	0.413
2-hr blood glucose in OGTT, mg/dL	116.90 ± 34.50	115.27 ± 33.05	0.055
≥ 140	1,052 (19.45)	428 (18.46)	0.313
Triglyceride, mg/dL	130.00 (96.0–183.0)	132.00 (99.0–183.0)	0.618
≥ 150	2,072 (38.31)	907 (39.15)	0.491
HDL-C, mg/dL	45.15 ± 10.19	44.77 ± 9.87	0.124
< 40 for males and < 50 for females	2,843 (52.57)	1,253 (54.06)	0.231
WBC, × 1,000/μL	6.50 ± 1.79	6.44 ± 1.78	0.189
γ-GTP, IU/L	20.00 (14.0–35.0)	20.00 (14.0–36.0)	0.335
Incident diabetes at the 8-year follow-up	460 (8.50)	180 (7.77)	0.280

Values are presented as mean ± standard deviation or median (interquartile range) or number (%).
BMI = body mass index, SBP = systolic blood pressure, DBP = diastolic blood pressure, HbA1c = hemoglobin A1c, OGTT = oral glucose tolerance test, HDL-C = high-density lipoprotein cholesterol, WBC = white blood cell, γ-GTP = γ-glutamyl transpeptidase.

had a low HDL-C level (52.6%). The mean FBG and HbA1c levels were 83.1 mg/dL and 5.6%, respectively. There were no significant differences between the derivation and validation datasets in terms of the baseline characteristics.

The modified FDRM included age (continuous variable; years), BMI (categorical variable; < 25.0, 25.0–29.9, or ≥ 30.0 kg/m²), HTN (≥ 140/90 mmHg or use of antihypertensive drugs), low HDL-C level (< 40 mg/dL for males and < 50 mg/dL for females), and elevated TG (≥ 150 mg/dL), FBG (100–125 mg/dL), and HbA1c (5.7–6.4%) levels. The stepwise selection method was used to develop an expanded clinical model including the white blood cell (WBC) count (continuous variable; × 1,000/μL) and log-transformed γ-glutamyl transpeptidase (γ-GTP) level (continuous variable; unit of raw value IU/L) to the modified FDRM. However, because the difference between the model adding WBC and γ-GTP (AUC = 0.831) and the model adding only γ-GTP (AUC = 0.830) was minimal (P = 0.474), only γ-GTP was included in the expanded clinical model. **Table 2** presents the beta coefficients of the models. The modified FDRM and expanded clinical models can be used to calculate the diabetes risk as follows:

Modified FDRM:
  Beta = −4.7926 + (age × 0.00964) + (HTN × 0.2724) + (high TG × 0.5917) + (hypo HDL-C × 0.05) + (high FBG × 2.0429) + (high HbA1c × 1.9041)

  if BMI 25.0–29.9 kg/m², + 0.2842; if BMI ≥ 30.0 kg/m², + 0.6765

Expanded clinical model:
  Beta = −6.7119 + (age × 0.0159) + (HTN × 0.2139) + (high TG × 0.3256) + (hypo HDL-C × 0.2596) + (high FBG × 1.8876) + (high HbA1c × 1.8972) + (log γ-GTP × 0.4906)

  if BMI 25.0–29.9 kg/m², + 0.2727; if BMI ≥ 30.0 kg/m², + 0.6705

Probability = 1/[1 + exp(−beta)]

Examples of the diabetes risk calculation are presented in the **Appendix 1**.

The relative importance of the features varied between the improved models (**Fig. 2**). HbA1c of 5.7–6.4% was the most important feature in both models, followed by an FBG level of

Table 2. Estimates and ORs for predictive models of incident diabetes based on the derivation dataset

Parameter	Modified Framingham Diabetes Risk Model		Expanded clinical model	
	Beta coefficients	OR (95% CI)	Beta coefficients	OR (95% CI)
Intercept	−4.7926		−6.7119	
Age, yr (continuous)	0.00964	1.01 (1.00–1.02)	0.0159	1.02 (1.00–1.03)
HTN, ≥ 140/90 mmHg or antihypertensive drug use	0.2724	1.31 (1.05–1.65)	0.2139	1.24 (0.99–1.56)
TG, ≥ 150 mg/dL	0.5917	1.81 (1.45–2.25)	0.3256	1.39 (1.09–1.75)
Hypo HDL-C, < 40 mg/dL for males and < 50 mg/dL for females	0.0500	1.05 (0.84–1.31)	0.2596	1.30 (1.03–1.64)
FBG, 100–125 mg/dL	2.0429	7.71 (5.70–10.43)	1.8876	6.60 (4.86–8.97)
HbA1c, 5.7–6.4%	1.9041	6.71 (5.20–8.66)	1.8972	6.67 (5.16–8.61)
BMI, kg/m²				
< 25.0	0	1.00	0	1.00
25.0–29.9	0.2842	1.33 (1.06–1.67)	0.2727	1.31 (1.04–1.66)
≥ 30.0	0.6765	1.97 (1.32–2.92)	0.6705	1.96 (1.31–2.91)
Log γ-GTP (continuous)			0.4906	1.63 (1.42–1.89)

OR = odds ratio, 95% CI = 95% confidence interval, HTN = hypertension, TG = triglyceride, HDL-C = high-density lipoprotein cholesterol, FBG = fasting blood glucose, HbA1c = hemoglobin A1c, BMI = body mass index, γ-GTP = γ-glutamyl transpeptidase.
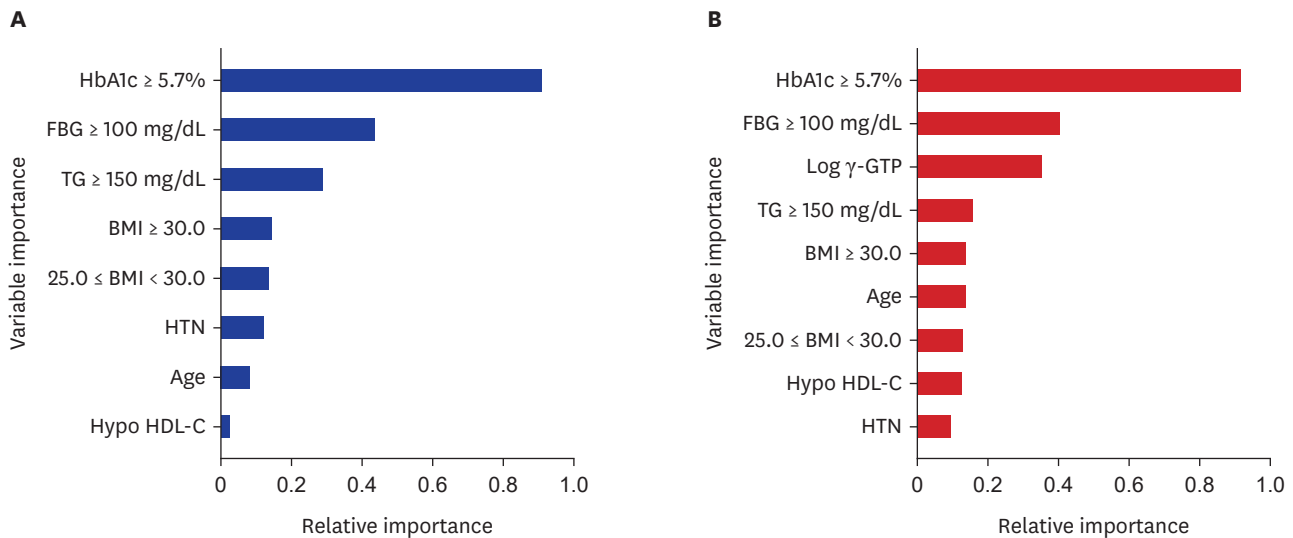
Fig. 2. Relative importance of features included in the improved predictive models for diabetes. (**A**) Modified Framingham Diabetes Risk Model and (**B**) expanded clinical model.
HTN was defied as blood pressure ≥ 140/90 mmHg or the use of antihypertensive drugs. Due to the skewed distribution, γ-GTP values were log-transformed.
Hypo HDL-C was defined as HDL-C < 40 mg/dL for males and < 50 mg/dL for females. Age, log-transformed γ-GTP, and WBC were included as numerical variables; all of the other variables were binary.
HbA1c = hemoglobin A1c, FBG = fasting blood glucose, TG = triglyceride, BMI = body mass index, HTN = hypertension, HDL-C = high-density lipoprotein cholesterol, WBC = white blood cell, γ-GTP = γ-glutamyl transpeptidase.

100–125 mg/dL. In the expanded clinical model, log-transformed γ-GTP ranked third in terms of importance and HTN was the least important variable. In the modified FDRM, a TG level ≥ 150 mg/dL ranked third and a low HDL-C level was the least important variable.

Fig. 3 compares the AUC values for the derivation and validation datasets among the predictive models. In the derivation dataset, the FDRM had the lowest AUC value (0.710, 95% CI, 0.684–0.736) and the expanded clinical model had the highest (0.830, 95% CI, 0.811–0.849); that of the modified FDRM was 0.820 (95% CI, 0.799–0.841). Both modified models had significantly higher AUC values than the FDRM (all $P < 0.001$), and the expanded clinical model had higher AUC values than the modified FDRM ($P = 0.004$). The accuracy of the expanded clinical model was 92.02% (95% CI, 91.27–92.73) and that of the modified FDRM was 91.93% (95% CI, 91.17–92.65). The accuracy of the FDRM was not calculated due to its low predictive value. When 10-fold cross-validation was applied, the AUCs of the modified FDRM and expanded clinical model were 0.813 and 0.825, respectively. The variance inflation factor of both models was < 1.5. We evaluated the interaction terms between the features of the predictive model. Although TG (≥ 150 mg/dL) and HbA1c (5.7–6.4%) showed a significant interaction, the interaction did not improve model performance (0.820 and 0.831 for the modified FDRM and expanded clinical model, respectively). Therefore, these features were not included in the model.

In the internal validation dataset, the discrimination abilities of the expanded clinical model and the modified FDRM were acceptable (0.822, 95% CI, 0.792–0.852 and 0.817, 95% CI, 0.785–0.850, respectively), and there was no significant difference between the models ($P = 0.602$). The accuracies of the expanded clinical model and the modified FDRM were similar to that in the derivation dataset (92.34%, 95% CI, 91.18–93.39 and 92.51%, 95% CI, 91.36–93.55, respectively; **Fig. 3**). In the temporal validation dataset, the γ-GTP data were not collected at the fourth follow-up, so data from the third follow-up were used. The AUC values
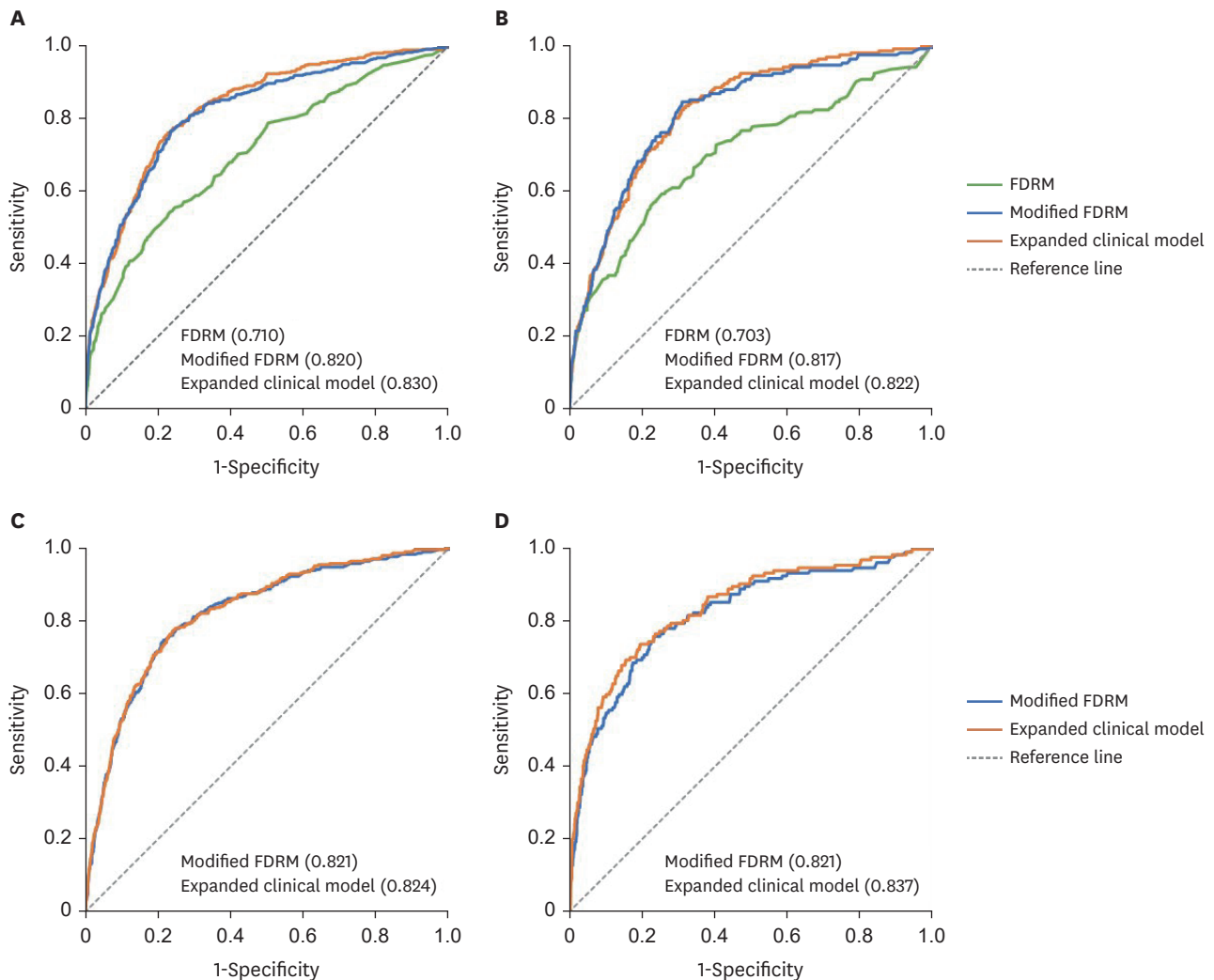
Fig. 3. Validation of the FDRM for predicting diabetes and comparison with the modified FDRM and expanded clinical model. (**A**) Derivation data, (**B**) internal validation data, (**C**) temporal validation data, and (**D**) external validation data.
FDRM = Framingham Diabetes Risk Model.

were 0.824 (95% CI, 0.800–0.848) and 0.821 (95% CI, 0.796–0.846) for the expanded clinical model and the modified FDRM (comparison between two models, P = 0.345), respectively. The accuracies of both the expanded clinical model and the modified FDRM were greater than 88% (88.41%, 95% CI, 87.28–89.47 and 88.32%, 95% CI, 87.29–89.28, respectively). In the external validation dataset, the AUCs of the expanded clinical model and the modified FDRM were 0.837 (95% CI, 0.800–0.875) and 0.821 (95% CI, 0.781–0.860), respectively. The accuracies of the expanded clinical model and modified FDRM were 92.51% (95% CI, 91.29–93.62) and 92.17% (95% CI, 90.92–93.30), respectively (**Fig. 3**). Temporal and external-validated data for participant characteristics are presented in **Supplementary Table 2**. They had higher average age, FBG, and lower TG than the participants in the derivation data and internal validation data.

Even when HTN according to the ACC/AHA standards was applied to the two modified models, the model performance was slightly lower, but generally showed a similar level (**Supplementary Table 3**). The cumulative incidence of incident diabetes, defined by adding

a 2-hour blood glucose level ≥ 200 mg/dL in the oral glucose tolerance test, was 13.53% (308/2,276) in internal validation data and 13.02% (490/3,764) in temporal validation, respectively. Overall, both modified models showed predictive power levels above 0.7, which were acceptable (**Supplementary Table 4**). The model performance of both modified models for men and women was above 0.8, with internal and temporal validation data being higher for women and external validation data being higher for men. However, the 95% CI were found to overlap. Although the overall model performance of both modified models was reduced in subjects with normoglycemia or prediabetes, model performance in prediabetic subjects was higher than that in normoglycemic subjects (**Supplementary Table 5**).

DISCUSSION

The FDRM showed acceptable predictive performance for incident diabetes in non-diabetic Korean individuals (AUC, 0.71); the predictive performance was improved by modifying the included factors (AUC, 0.82 for the modified FDRM). Although the addition of clinical features to the modified FDRM provided further improvement (AUC, 0.83), it was not maintained in the validation data. Furthermore, although the characteristics of the population-based cohort (e.g., KoGES-CAVAS) and the temporal validation dataset were different from those of the derived data, the modified FDRM showed good performance in each dataset (AUC, 0.82 and 0.82, respectively). This model performed equally well in both men and women. Even when diabetes prediction probability was calculated by defining HTN according to ACC/AHA, the overall model performance did not change significantly.

The FDRM has been validated for use in Taiwanese,[16] German,[17] Swedish,[18] American,[19] and Canadian populations,[11] with only some studies reporting acceptable performance. The different results among previous studies may have been due to the heterogeneity in the distribution of predictors and disease.[11,19] Indeed, in the present study, the participants were thinner (24.4 vs. 27.1 kg/m²) and had higher proportions of low HDL-C level (52.6% vs. 36.9%) and incident diabetes (8.5% vs. 5.1%) than in the Framingham offspring study.[5] Previous studies, including the Framingham offspring study, diagnosed diabetes on the basis of the FBG level, 2-hour postprandial blood glucose level, or use of antidiabetic drugs.[5,18,20] The American Diabetes Association recommends HbA1c as a substitute for FBG for the diagnosis of diabetes, and as a reliable prognostic marker for diabetes.[6,21] HbA1c reflects the mean plasma glucose level over the previous 8–12 weeks.[22] In the present study, HbA1c had the highest importance among the features included in the prediction model. Changes in diagnostic criteria may affect the performance of the predictive model. Previous studies of the predictive ability of clinical models based on 2-hour glucose and FBG levels have shown conflicting results.[23,24] The diagnosis guidelines for diabetes recommend a 2-hour glucose-tolerance test, but it is not mandatory for screening. Studies that developed diabetes prediction models using health checkup data generally did not consider 2-hour glucose when defining diabetes.[25,26] However, in line with previous studies,[27,28] we observed differences in prevalence depending on the definition of diabetes. Also, one meta-study suggested lowering the cutoff values for FBG and HbA1c in relation to diabetes diagnosis.[28] Therefore, prior to developing a diabetes prediction model, it seems necessary to resolve this discrepancy.

Various prediction models for diabetes have been developed over the last 20 years.[4] Age, family history of diabetes, HTN, BMI or obesity, waist circumference, and sex have frequently been included in prediction models of diabetes.[10] To reduce the cost and the time required to

predict diabetes risk, Lindstrom and Tuomilehto[20] proposed the Finnish Diabetes Risk Score based on lifestyle features, rather than laboratory tests, to identify high-risk individuals; the score had an AUC of 0.85 for the prediction of diabetes. Stern et al.[23] developed a predictive model based on the components of metabolic syndrome for Mexican-American and non-Hispanic whites. The model which consisted of metabolic features, had higher discriminative ability than the 2-hour oral glucose tolerance test alone for high-risk individuals.[23] In the present study, the discrimination abilities of 2-hour glucose, HbA1c (5.7–6.4%), and FBG levels (100–125 mg/dL) were 0.78, 0.74, and 0.61, respectively (**Supplementary Table 1**). Our findings also showed that the modified FDRM had a higher 8-year prediction ability for diabetes than univariate models based on 2-hour glucose, HbA1c, and FBG levels. Several studies have found that metabolic features are reasonable and useful for the prediction of diabetes. However, as with the Framingham offspring study,[5] making models more complex by adding clinical features did not improve their performance. We selected γ-GTP and WBC as potential risk factors of diabetes. Previous studies have identified an association between chronic subclinical inflammation and insulin resistance. In addition, several studies found an association between an elevated WBC count and diabetes.[29,30] In a study of Taiwanese individuals, a predictive model based on age, BMI, WBC count, TG, HDL-C, and FPG levels was developed and had an AUC of 0.70.[16] However, it did not appear to primarily contribute to the performance of the diabetes prediction model, so it was not included in the expanded clinical model. Liver-related biomarkers are associated with diabetes, and γ-GTP is an early and sensitive marker of inflammation and oxidative stress.[31] In addition, it has been suggested that γ-GTP is an independent predictor of diabetes, HTN, metabolic syndrome, and coronary artery disease.[32] A study by Lee et al.[33] from the same data source as this study also found that higher γ-GTP was associated with the development of diabetes. Another Korean study using health examination data also found that γ-GTP contributed to the prediction of diabetes.[25] The role of oxidative stress in the development of diabetes-related complications is unclear; some studies found associations with peripheral neuropathy and retinopathy.[34,35] Although the Framingham offspring study developed a simple predictive model, the use of a parental history of diabetes is challenging. Parental history of diabetes is asked in routine clinical settings, but its availability in research appears to be limited. A Canadian study that analyzed electronic medical records as primary care data evaluated the FDRM after excluding a parental history of diabetes (due to a lack of data),[11] whereas another study used family history instead of a parental history of diabetes.[19] In external validation data, data on parental history of diabetes was missing in 26.2% of subjects. In addition, in our study, no significant contribution of parental history of diabetes was observed in predicting diabetes. Also, like previous studies,[5,25] the contribution of sex to diabetes prediction was not found to be significant. Instead, the modified FDRM showed a predictive power of over 0.8 for both sexes.

Certain factors should be considered when interpreting our results. First, we evaluated the validity of the modified FDRM in various data settings. Further studies are needed to determine the validity of the model in other populations. Second, misclassification bias may have occurred due to measurement error.

Our results have significant implications because we evaluated the usefulness of the FDRM and proposed a modified FDRM in consideration of utilization in the clinical setting. In addition, the validity of the modified model was assessed using various datasets in terms of temporal validation and external validation. In total, 37.9% and 36.8% of the population had prediabetes (HbA1c of 5.7–6.4%) in the derivation and validation datasets, respectively;

this variable made the largest contribution to the prediction of incident diabetes (**Fig. 2**). Therefore, for individuals with prediabetes, interventions to prevent or delay the onset of diabetes are required. In primary care, high-risk individuals should be targeted for counseling and interventions to prevent diabetes.

Taken together, our results show that the predictive performance of the FDRM was acceptable in a Korean general population, although lower than that in the Framingham offspring study. The discrimination ability of the FDRM was improved when the redefined features based on recent clinical criteria were added. However, adding clinical features to the modified FDRM did not improve model performance. Identification of high-risk individuals and early intervention may reduce the diabetes burden. Further studies are needed to confirm the validity of the diabetes-prediction models in various populations.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIALS

Supplementary Table 1
Univariate analysis of features associated with incident diabetesa during the 8-year follow-up

Supplementary Table 2
Basic characteristics of the participants in external and temporal validation data

Supplementary Table 3
Sensitivity analysis for predication models

Supplementary Table 4
Sensitivity analysis for predication models according to diabetes definition

Supplementary Table 5
Model performance for predicting incident diabetes according to subgroups

REFERENCES

1. World Health Organization, Chronic Respiratory Diseases and Arthritis Team. Screening for type 2 diabetes: report of a World Health Organization and International Diabetes Federation meeting. https://apps.who.int/iris/handle/10665/68614. Updated 2003. Accessed March 1, 2023.

2. International Diabetes Federation. *IDF Diabetes Atlas, 10th ed.* Brussels, Belgium: International Diabetes Federation; 2021.

3. Herman WH, Ye W, Griffin SJ, Simmons RK, Davies MJ, Khunti K, et al. Early detection and treatment of type 2 diabetes reduce cardiovascular morbidity and mortality: a simulation of the results of the Anglo-Danish-Dutch Study of Intensive Treatment in People with Screen-Detected Diabetes in Primary Care (ADDITION-Europe). *Diabetes Care* 2015;38(8):1449-55. **PUBMED | CROSSREF**

4. Fregoso-Aparicio L, Noguez J, Montesinos L, García-García JA. Machine learning and deep learning predictive models for type 2 diabetes: a systematic review. *Diabetol Metab Syndr* 2021;13(1):148. **PUBMED | CROSSREF**

5. Wilson PW, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB Sr. Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. *Arch Intern Med* 2007;167(10):1068-74.
PUBMED | **CROSSREF**

6. World Health Organization. Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus. https://apps.who.int/iris/bitstream/handle/10665/70523/WHO_NMH_CHP_CPM_11.1_eng.pdf. Updated 2011. Accessed March 1, 2023.

7. Flack JM, Adekola B. Blood pressure and the new ACC/AHA hypertension guidelines. *Trends Cardiovasc Med* 2020;30(3):160-4. **PUBMED** | **CROSSREF**

8. Kim HL, Lee EM, Ahn SY, Kim KI, Kim HC, Kim JH, et al. The 2022 focused update of the 2018 Korean Hypertension Society guidelines for the management of hypertension. *Clin Hypertens* 2023;29(1):11.
PUBMED | **CROSSREF**

9. Kim Y, Han BG; KoGES group. Cohort profile: the Korean Genome and Epidemiology Study (KoGES) consortium. *Int J Epidemiol* 2017;46(2):e20. **PUBMED** | **CROSSREF**

10. Collins GS, Mallett S, Omar O, Yu LM. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. *BMC Med* 2011;9(1):103. **PUBMED** | **CROSSREF**

11. Mashayekhi M, Prescod F, Shah B, Dong L, Keshavjee K, Guergachi A. Evaluating the performance of the Framingham Diabetes Risk Scoring Model in Canadian electronic medical records. *Can J Diabetes* 2015;39(2):152-6. **PUBMED** | **CROSSREF**

12. Choi HS, Lee SW, Kim JT, Lee HK. The association between pulmonary functions and incident diabetes: longitudinal analysis from the Ansung cohort in Korea. *Diabetes Metab J* 2020;44(5):699-710. **PUBMED** | **CROSSREF**

13. Moon S, Jang JY, Kim Y, Oh CM. Development and validation of a new diabetes index for the risk classification of present and new-onset diabetes: multicohort study. *Sci Rep* 2021;11(1):15748. **PUBMED** | **CROSSREF**

14. Kim KN, Oh SY, Hong YC. Associations of serum calcium levels and dietary calcium intake with incident type 2 diabetes over 10 years: the Korean Genome and Epidemiology Study (KoGES). *Diabetol Metab Syndr* 2018;10(1):50. **PUBMED** | **CROSSREF**

15. Park S, Kim C, Wu X. Development and validation of an insulin resistance predicting model using a machine-learning approach in a population-based cohort in Korea. *Diagnostics (Basel)* 2022;12(1):212.
PUBMED | **CROSSREF**

16. Chien K, Cai T, Hsu H, Su T, Chang W, Chen M, et al. A prediction model for type 2 diabetes risk among Chinese people. *Diabetologia* 2009;52(3):443-50. **PUBMED** | **CROSSREF**

17. Li J, Bornstein SR, Landgraf R, Schwarz PE. Validation of a simple clinical diabetes prediction model in a middle-aged, white, German population. *Arch Intern Med* 2007;167(22):2528-9. **PUBMED** | **CROSSREF**

18. Lyssenko V, Jonsson A, Almgren P, Pulizzi N, Isomaa B, Tuomi T, et al. Clinical risk factors, DNA variants, and the development of type 2 diabetes. *N Engl J Med* 2008;359(21):2220-32. **PUBMED** | **CROSSREF**

19. Nichols GA, Brown JB. Validating the Framingham Offspring Study equations for predicting incident diabetes mellitus. *Am J Manag Care* 2008;14(9):574-80. **PUBMED**

20. Lindström J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. *Diabetes Care* 2003;26(3):725-31. **PUBMED** | **CROSSREF**

21. Sherwani SI, Khan HA, Ekhzaimy A, Masood A, Sakharkar MK. Significance of HbA1c test in diagnosis and prognosis of diabetic patients. *Biomark Insights* 2016;11:95-104. **PUBMED** | **CROSSREF**

22. Nathan DM, Turgeon H, Regan S. Relationship between glycated haemoglobin levels and mean glucose levels over time. *Diabetologia* 2007;50(11):2239-44. **PUBMED** | **CROSSREF**

23. Stern MP, Williams K, Haffner SM. Identification of persons at high risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? *Ann Intern Med* 2002;136(8):575-81. **PUBMED** | **CROSSREF**

24. McNeely MJ, Boyko EJ, Leonetti DL, Kahn SE, Fujimoto WY. Comparison of a clinical model, the oral glucose tolerance test, and fasting glucose for prediction of type 2 diabetes risk in Japanese Americans. *Diabetes Care* 2003;26(3):758-63. **PUBMED** | **CROSSREF**

25. Shin J, Kim J, Lee C, Yoon JY, Kim S, Song S, et al. Development of various diabetes prediction models using machine learning techniques. *Diabetes Metab J* 2022;46(4):650-7. **PUBMED** | **CROSSREF**

26. Jeong YW, Jung Y, Jeong H, Huh JH, Sung KC, Shin JH, et al. Prediction model for hypertension and diabetes mellitus using Korean public health examination data (2002–2017). *Diagnostics (Basel)* 2022;12(8):1967. **PUBMED** | **CROSSREF**

27. Jeon JY, Ko SH, Kwon HS, Kim NH, Kim JH, Kim CS, et al. Prevalence of diabetes and prediabetes according to fasting plasma glucose and HbA1c. *Diabetes Metab J* 2013;37(5):349-57. **PUBMED** | **CROSSREF**

28. Kaur G, Lakshmi PV, Rastogi A, Bhansali A, Jain S, Teerawattananon Y, et al. Diagnostic accuracy of tests for type 2 diabetes and prediabetes: a systematic review and meta-analysis. *PLoS One* 2020;15(11):e0242415. **PUBMED** | **CROSSREF**

29. Kashima S, Inoue K, Matsumoto M, Akimoto K. White blood cell count and C-reactive protein independently predicted incident diabetes: Yuport Medical Checkup Center Study. *Endocr Res* 2019;44(4):127-37. **PUBMED** | **CROSSREF**

30. Chen JY, Chen YH, Lee YC, Tsou MT. The association between white blood cell count and insulin resistance in community-dwelling middle-aged and older populations in Taiwan: a community-based cross-sectional study. *Front Med (Lausanne)* 2022;9:813222. **PUBMED** | **CROSSREF**

31. Onur S, Niklowitz P, Jacobs G, Nöthlings U, Lieb W, Menke T, et al. Ubiquinol reduces gamma glutamyltransferase as a marker of oxidative stress in humans. *BMC Res Notes* 2014;7(1):427. **PUBMED** | **CROSSREF**

32. Onat A, Can G, Örnek E, Çiçek G, Ayhan E, Doğan Y. Serum γ-glutamyltransferase: independent predictor of risk of diabetes, hypertension, metabolic syndrome, and coronary disease. *Obesity (Silver Spring)* 2012;20(4):842-8. **PUBMED** | **CROSSREF**

33. Lee JH, Lee HS, Lee YJ. Serum γ-glutamyltransferase as an independent predictor for incident type 2 diabetes in middle-aged and older adults: findings from the KoGES over 12 years of follow-up. *Nutr Metab Cardiovasc Dis* 2020;30(9):1484-91. **PUBMED** | **CROSSREF**

34. Cho HC. The association between serum GGT concentration and diabetic peripheral polyneuropathy in type 2 diabetic patients. *Korean Diabetes J* 2010;34(2):111-8. **PUBMED** | **CROSSREF**

35. Valizadeh N, Mohammadi R, Mehdizadeh A, Motarjemizadeh Q, Khalkhali HR. Evaluation of serum gamma glutamyl transferase levels in diabetic patients with and without retinopathy. *Shiraz E Med J* 2018;19(7):e64073. **CROSSREF**

Appendix 1. Example diabetes risk calculation using the modified FDRM and expanded clinical model

Parameter	Modified FDRM		Expanded clinical model	
	Beta coefficients	Case 1	Beta coefficients	Case 1
Intercept	−4.7926		−6.7119	
Age, yr (continuous)	0.00964	50	0.0159	50
HTN, ≥ 140/90 mmHg or antihypertensive drug use	0.2724	Present (0)	0.2139	Present (0)
TG, ≥ 150 mg/dL	0.5917	Absent (0)	0.3256	Absent (0)
Hypo HDL-C, < 40 mg/dL for males and < 50 mg/dL for females	0.0500	Present (1)	0.2596	Present (1)
FBG, 100–125 mg/dL	2.0429	Absent (0)	1.8876	Absent (0)
HbA1c, 5.7–6.4%	1.9041	Present (1)	1.8972	Present (1)
BMI, kg/m²				
< 25.0	0	Absent (0)	0	Absent (0)
25.0–29.9	0.2842	Present (1)	0.2727	Present (1)
≥ 30.0	0.6765	Absent (0)	0.6705	Absent (0)
log γ-GTP (continuous)			0.4906	2.398

FDRM = Framingham Diabetes Risk Model, HTN = hypertension, TG = triglyceride, HDL-C = high-density lipoprotein cholesterol, FBG = fasting blood glucose, HbA1c = hemoglobin A1C, BMI = body mass index, γ-GTP = γ-glutamyl transpeptidase.

The probability of incident diabetes of a 50-year-old female with a BMI of 25.5 kg/m², fasting blood glucose level of 99 mg/dL, HbA1c of 5.9%, TG of 91 mg/dL, HDL-C of 47 mg/dL, SBP/DBP of 118/89 mmHg, no anti-hypertensive drug use, and γ-GTP of 11 IU/L can be calculated as follows:

Modified FDRM:
Beta = −4.7926 + [age (50) × 0.00964] + [HTN (0) × 0.2724] + [high TG (0) × 0.5917] + [hypo HDL-C (1) × 0.0500] + [high FBG (0) × 2.0429] + [high HbA1c (1) × 1.9041] + [25.0 ≤ BMI < 30.0 (1) × 0.2842] = −2.0723
Probability = $1/\{1+\exp[-(-2.0723)]\}$ = 0.112, 11.2%

Expanded clinical model:
Beta = −6.7119 + [age (50) × 0.0159] + [HTN (0) × 0.2139] + [high TG (0) × 0.3256] + [hypo HDL-C (1) × 0.2596] + [high FBG (0) × 1.8876] + [high HbA1c (1) × 1.8972] + [log γ-GTP (2.3979) × 0.4906] + [25.0 ≤ BMI < 30.0 (1) × 0.2727] = −2.3110
Probability = $1/\{1+\exp[-(-2.3110)]\}$ = 0.090, 9.0%

For this case, the probabilities of developing diabetes according to the modified FDRM and expanded clinical model were estimated to be 11.2% and 9.0%, respectively.