Original Article

Non-invasive Risk Prediction Models in Identifying Undiagnosed Type 2 Diabetes or Predicting Future Incident Cases in the Iranian Population

Mojtaba Lotfaliany, MD[1,2]; Farzad Hadaegh, MD[2]; Samaneh Asgari, PhD[2]; Mohammad Ali Mansournia, PhD[3]; Fereidoun Azizi, MD[4]; Brian Oldenburg, PhD[1]; Davood Khalili, PhD[2,5*]

[1]School of Population and Global Health, University of Melbourne, Australia
[2]Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[3]Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran
[4]Endocrine Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[5]Department of Biostatistics and Epidemiology, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Abstract
Background: Iran needs pragmatic screening methods for identifying those with undiagnosed type 2 diabetes or at high risk of developing it. The aim of this study was to assess performance of three non-invasive risk prediction models, i.e. the Finnish Diabetes Risk Score (FINDRISC), the Australian Type 2 Diabetes Risk Assessment Tool (AUSDRISK), and the American Diabetes Association Risk Score (ADA), for identifying those with undiagnosed type 2 diabetes (prevalent type 2 diabetes at baseline without any treatment) or those who would develop type 2 diabetes within 5 years of follow-up
Methods: 3467 participants aged ≥30 years without treated type 2 diabetes in the Tehran Lipid and Glucose Study (TLGS) were included in this study. The discrimination power of models was assessed by area under the curve (AUC), their calibrations were assessed by calibration plots and Hosmer–Lemeshow test, and their net benefits were assessed by decision curves.
Results: 430 participants had undiagnosed type 2 diabetes at baseline and 203 developed type 2 diabetes during 5 years of follow-up. AUSDRISK had the highest AUC (0.77) as compared to FINDRISC (0.75; *P* value: 0.014), and the ADA model (0.73; *P* value: <0.001). The original model for AUSDRISK and calibrated versions of FINDRISC and ADA models had acceptable calibration (Hosmer–Lemeshow chi-square <20) and these models were clinically useful in a wide range of risk thresholds as their net benefit was higher than no-screening scenarios.
Conclusion: The original AUSDRISK model and recalibrated models for FINDRISC and ADA are valid and effective tools for identifying those with undiagnosed or 5-year incident type 2 diabetes in Iran.
Keywords: High risk individuals, Non-invasive risk prediction models, Screening methods, Type 2 diabetes
Cite this article as: Lotfaliany M, Hadaegh F, Asgari S, Mansournia MA, Azizi F, Oldenburg B, et al. Non-invasive risk prediction models in identifying undiagnosed type 2 diabetes or predicting future incident cases in the Iranian population. Arch Iran Med. 2019;22(3):116–124.

Introduction

In 2017, Iran ranked second for the highest number of people with type 2 diabetes in the Middle East region, with 8.9% of the adult population having the disease (5 million people), about 35% of whom are undiagnosed.[1] Moreover, in our previous study in 2008, we showed that Iranians have high risk for type 2 diabetes with 12.3% of the Iranian population having impaired fasting glucose and 11.4% of them having impaired glucose tolerance.[2] There is strong evidence showing that lifestyle or pharmacological interventions can prevent type 2 diabetes in individuals at high risk for type 2 diabetes[3-6] both of which are most effective in short-term (<5 years).[7] Moreover, early diagnosis of those with undiagnosed type

2 diabetes can significantly prevent the complications of type 2 diabetes, thereby reducing its burden.[8,9] Given the high burden of type 2 diabetes in Iran and the availability of pragmatic interventions, there is the need for pragmatic screening methods for identifying those at high risk of type 2 diabetes or with undiagnosed type 2 diabetes in the Iranian population.

Current guidelines recommend non-invasive risk prediction models for screening undiagnosed type 2 diabetes and identifying those at high risk of for type 2 diabetes.[10-12] In most real-world settings like the national type 2 diabetes prevention program in Finland,[13] risk prediction models are used among those without known type 2 diabetes (i.e. those without type 2 diabetes and

*Corresponding Author: Davood Khalili, MD, MPH, PhD; Ass. Professor of Epidemiology, Prevention of Metabolic Disorders Research Center, & Head of Department of Biostatistics and Epidemiology, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Postal Address: No. 23, Parvaneh Street, Velenjak, Tehran, Iran. Phone: +98-21- 22432500; Fax: +98-21- 22416264; Email: dkhalili@endocrine.ac.ir

those with undiagnosed type 2 diabetes) to identify those at high risk. In the next step, high risk individuals undergo blood tests and receive an appropriate intervention based on results of blood tests (i.e. medication for type 2 diabetes for those with undiagnosed type 2 diabetes or preventive interventions for those with high risk of developing it).

A number of non-invasive and easily administrable diabetes risk prediction models have been developed for detecting those with undiagnosed type 2 diabetes and/ or high risk individuals for type 2 diabetes in different settings.[14,15] These have been used in several clinical trials for prevention of type 2 diabetes to identify high risk individuals, with the FINDRISC (Finnish Diabetes Risk Score),[16] AUSDRISK (Australian Type 2 Diabetes Risk Assessment Tool),[17] and ADA (American Diabetes Association Risk Score)[18] risk prediction models being the most common risk prediction models used[3,6]; several studies also used these three risk prediction models to identify those with undiagnosed type 2 diabetes[19-23] and previous systematic reviews also show that these three risk scores have the potential to be used in clinic settings.[24]

Despite these three risk prediction model yielding very promising findings in studies conducted in high-income countries, there are only a few studies in low- and middle-income countries that have assessed the performance of these risk prediction models in identifying those with undiagnosed type 2 diabetes and/or at high risk of developing it.[25] Therefore, it remains unknown whether these three risk prediction models can be used in screening. This study aims to assess the discrimination power, calibration, and net benefit of the FINDRISC, AUSDRISK, and ADA risk prediction models for identifying individuals at high risk of type 2 diabetes (5-year incident type 2 diabetes) or with undiagnosed type 2 diabetes (prevalent type 2 diabetes at baseline without any treatment) who may benefit from interventions for type 2 diabetes in a community-representative sample of people living in Tehran, the capital city of Iran. In the secondary analyses, we assessed performance of these three risk prediction models for each outcome of 5-year incidence of type 2 diabetes and undiagnosed type 2 diabetes separately.

Materials and Methods

Study Subjects

The Tehran Lipid and Glucose Study (TLGS) is a longitudinal study being conducted in a community-representative sample of Tehran, capital of Iran, with the aim of determining the prevalence and incidence of non-communicable disease and related risk factors. Details of TLGS have been reported previously.[26] Briefly, data collection of 15,005 individuals was initiated in 1999-2001 (phase 1) and all participants were re-examined triennially (phases 2–5).[26] Of 15 005 participants, 5,630 individuals were assigned to a lifestyle modification intervention,[26] which significantly reduced the incidence

of type 2 diabetes in the intervention group.[27] Therefore, in this study, only participants in the control arm and those aged ≥30 years were included (n = 4908). Moreover, 257 participants with drug-treated type 2 diabetes and 44 participants with pregnancy were excluded. Of 4607 eligible participants, 271 participants with no information on type 2 diabetes at baseline, 69 participants with no data on type 2 diabetes on follow-up examinations, 645 participants with no follow-up, 155 participants will total follow-up <5 years were excluded, leaving 3467 participants for the main analyses. In secondary analyses, to assess the performance of the models in identifying those with 5-year incident type 2 diabetes, we further excluded 430 participants with undiagnosed type 2 diabetes. Since for assessing performance of the models in identifying those with undiagnosed type 2 diabetes, the follow-up data was not of interest, 4336 participants with available data at baseline were included in the analysis.

Clinical and Laboratory Measurements

Details of clinical and laboratory measurements have been reported elsewhere.[26] The workflow was designed according to the World Health Organization (WHO) stepwise approach to non-communicable disease surveillance.[28] Participants were interviewed to obtain demographics and past medical history by completing a 110-item standardized and validated questionnaire. Physical activity level was assessed with the Lipid Research Clinic questionnaire.[29] Anthropometric measurements were taken with shoes removed and the participants wearing light clothing. Weight and height were measured according to the standard protocol. Waist circumference (WC) was measured at the level of the umbilicus. For measuring blood pressure, the participants remained seated for 15 minutes, then a qualified physician measured blood pressure twice after one more measurement for determining peak inflation level, using a standard and calibrated mercury sphygmomanometer. Blood samples were drawn between 7:00 and 9:00 AM into vacutainer tubes from all study participants after 12–14 hours overnight fasting. All blood analyses were done at the TLGS research laboratory on the day of blood collection. For OGTT, 75 g anhydrous glucose was administered orally. Fasting plasma glucose (FPG) and 2 hour-post-challenge plasma glucose (2h-PG) were measured using an enzymatic colourimetric method with glucose oxidase; inter- and intra-assay coefficients of variation at baseline and follow-up phases were both less than 2.3%. High-density lipoprotein cholesterol (HDL-C) was measured after precipitation of the apolipoprotein B containing lipoproteins with phosphotungstic acid. Triglycerides were assayed using glycerol phosphate oxidase.

Definition of Terms

The outcome was undiagnosed type 2 diabetes at baseline

or incidence of type 2 diabetes within the first 5 years of follow-up. Type 2 diabetes was ascertained in each phase of TLGS among participants who had FPG ≥7.0 mmol/L or postprandial plasma glucose ≥11.1 mmol/L and/or were taking glucose-lowering medication. Undiagnosed type 2 diabetes at baseline was defined as having FPG ≥ 7.0 mmol/L or postprandial plasma glucose ≥11.1 without treatment of the participants, based on self-report. The event date was considered as the half-time between the first date that the type 2 diabetes was diagnosed and the last known disease-free date. Censor date was defined as the last follow-up date in those without incident type 2 diabetes. Based on the event date, censor date, and type 2 diabetes status, participants were divided to 2 groups: (1) those with 5-year incident type 2 diabetes including those with type 2 diabetes at baseline and (2) those without 5-year incident type 2 diabetes.

The risk of developing outcome was estimated for each individual based on the risk prediction models at the baseline examination. The predicted risk was compared to observed risk to compare the performance of the risk prediction models. Body mass index (BMI) was calculated by dividing the weight in kilograms by the square of height in meters. Current smoking was ascertained in those participants who smoked cigarettes at least once a day or who smoked cigarettes occasionally. The family history of diabetes was defined as having at least one parent or sibling with diabetes.

Brief Description of Risk Prediction Models

The FINDRISC model was developed in a Finn population sample of 35–64-year-olds with no antidiabetic drug treatment at baseline (the study had 2 parts initiated in 1987 and 1992) who were followed for 10 years.[16] FINDRISC uses age, BMI, WC, physical activity, daily consumption of fruits, berries, or vegetables, and the history of antihypertensive drug treatment and history of high blood glucose to predict drug-treated diabetes[16]; there is also a concise version for FINDRISC model that did not use physical activity, and daily consumption of fruits, berries, or vegetables.[16] Due to lack of data for daily consumption of vegetables, berries, or fruits in the baseline examination of TLGS, the concise version of the FINDRISC was used in the current study.[16] AUSDRISK was developed in an Australian population aged ≥25 years without physician-diagnosed diabetes at baseline (year: 1999), who were followed for 5 years[17]; it uses age, sex, ethnicity, parental history of diabetes, history of high blood glucose level, use of antihypertensive medications, smoking, physical inactivity, and WC to predict the incidence of type 2 diabetes based on fasting plasma glucose and 2-hour plasma glucose levels.[17] The ADA risk prediction model developed in an American population aged ≥20 years without self-reported diagnosed type 2 diabetes (year: 1999) to detect undiagnosed type 2 diabetes

based on age, sex, family history of diabetes, history of hypertension, obesity, and physical activity.[18]

Statistical Analysis

Baseline characteristics were summarized as mean (standard deviation: SD) values for continuous and frequencies (%) for categorical variables in those with and without type 2 diabetes. Since the blood level of triglycerides had a skewed distribution, it was summarized by the median (IQR). The baseline characteristics of participants with different outcomes (i.e. undiagnosed type 2 diabetes at baseline, no diabetes after 5 years of follow-up, and 5-year incident type 2 diabetes) were compared between by Student's t test for continuous variables, the Chi-square test for categorical variables, and the Mann-Whitney U statistic for skewed variables to highlight the differences between groups.

Observations for BMI, WC, physical activity, and smoking status were missing in 70 (2.0%), 71 (2.0%), 68 (2.0%), and 65 (1.9%) records, respectively. Single imputation was performed to impute missing values for variables with missing data. For imputing the missing values of BMI and WC linear regression and for imputing the missing values of physical activity and smoking status, logistic regression, was fitted using age, sex, education status, family history of diabetes, self-report of hypertension, hyperglycemia, drug history for hypertension, systolic and diastolic blood pressure, FPG, 2h-PG, HDL-C, total cholesterol, and triglycerides, and dyslipidemia as auxiliary variables. The imputation of missing values before development and/or validation of risk prediction models has been recommended before.[30] In the sensitivity analyses, we further imputed the missing values of the type 2 diabetes status in those eligible participants who were excluded from the analyses due to missing data for type 2 diabetes at baseline examination or after 5 years of follow up using similar methods for imputation (n = 1140).

A logistic regression was fitted with type 2 diabetes as the outcome and the linear predictor part of each model as an offset variable to calibrate the risk prediction models' intercept, developing a "calibrated-in-the-large" model for each risk prediction model.[31] Furthermore, a separate logistic regression was fitted with type 2 diabetes as the outcome and the linear predictor part of each model as the only predictor variable to derive a calibration slope and a new intercepts, developing a "recalibrated" model for each risk prediction model.[31]

To assess the discrimination power of the models, a receiver operating characteristic curve was plotted and the area under the curve (AUC) has been estimated for each risk prediction model. Calibration of the original, calibrated-in-the-large, and recalibrated models were assessed visually as well as using Hosmer-Lemeshow chi-square; a Hosmer-Lemeshow chi-square higher than 20 was defined as the clear evidence for lack of calibration.[32]

In line with guideline of the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis Initiative,[33] we plotted predicted outcome probabilities (x-axis) against observed outcomes (y-axis) using a LOWESS (locally weighted scatterplot smoothing) line.

Net benefit of a risk threshold (P) was calculated based on the number of true positives (TP) and false positives (FP) of model if the risk threshold was used as a cut-off for identifying those with the outcome as well as total number of the population (N), and odds of the risk threshold $[Odds_p = P/(1-P)]$ in the following formula[34]:

Net benefit = (TP – $Odds_p$ × FP)/N

The risk threshold is the threshold probability in which the expected benefit of the intervention is equal to the expected benefit of avoiding the intervention and can be assessed by the following formula[34]:

$Odds_p = P/(1-P) = Harm/Benefit$

The risk threshold in a setting should be derived considering costs and performance of screening method, effectiveness, costs, and adverse effects of the intervention as well as available resources for the screening and intervention to derive a harm-to-benefit ratio; it may therefore vary in different settings. To assess the range of risk thresholds in which the risk prediction models are useful, the net benefit of the models were plotted in the wide range of risk thresholds to draw decision curves. The net benefit of the risk prediction models were compared with the net benefits of two different scenarios without screening: treat-none and treat-all. The risk prediction model was useful in a risk threshold if the differences between net benefit of risk prediction model and no-screening methods were equal or higher than 0.01; cut-off of 0.01 for difference in net benefit for non-invasive or minimal invasive screening methods has been suggested previously and it is equivalent to a net number to treat/test of 100.[35] Analyses were repeated for each secondary outcome. All analyses were performed using Stata statistical software (version 14 SE).

Results

Of the 4607 eligible participants, 3467 participants (75%) were included in this study, of whom 430 (12.4%) had undiagnosed type 2 diabetes at baseline and 203 (5.8%) developed type 2 diabetes during 5 years of follow-up. Table 1 compares the baseline characteristics of participants. Participants aged 50 years on average and 55% were female and their average BMI was 27.6 kg/m². Participants who developed type 2 diabetes within 5 years of follow-up and those with undiagnosed type 2 diabetes had higher level of risk factors at baseline as compared to those with no type 2 diabetes at end of follow-up.

Figure 1 shows the receiver operating characteristic curves of the models. For the main outcome (i.e. undiagnosed type 2 diabetes at baseline or 5-year incident type 2 diabetes), AUSDRISK had the highest discrimination power (AUC = 0.77) as compared to FINDRISC (AUC = 0.75; P value = 0.030) and ADA risk prediction (AUC = 0.73; P value

Table 1. Baseline Characteristics of Participants

Characteristics	Undiagnosed Type 2 Diabetes at Baseline (n = 430)	No Type 2 Diabetes at Baseline (n = 3037)		P Value[a]
		5-Year Incident Type 2 Diabetes (n = 203)	No 5-Year Incident Type 2 Diabetes (n = 2834)	
Age	54.73 (11.70)	49.93 (11.15)	45.57 (11.14)	<0.001
Sex (%)				
Female	235 (54.7%)	121 (59.6%)	1564 (55.2%)	0.450
Education (%)				
Less than 6 years	248 (57.7%)	112 (55.4%)	1022 (36.1%)	<0.001
6-12 years	148 (34.4%)	75 (37.1%)	1432 (50.6%)	
More than 12 years	34 (7.9%)	15 (7.4%)	375 (13.3%)	
Physically inactive (%)	312 (75.5%)	143 (72.2%)	2074 (74.4%)	0.680
Current smoking (%)	49 (11.8%)	27 (13.6%)	429 (15.4%)	0.150
Weight (kg)	74.78 (11.87)	75.53 (12.85)	71.51 (12.09)	<0.001
BMI (kg/m²)	29.20 (4.22)	29.39 (4.74)	27.19 (4.34)	<0.001
WC (cm)	96.61 (10.41)	95.63 (10.43)	89.05 (10.93)	<0.001
Systolic blood pressure (mm Hg)	134.94 (24.21)	129.37 (20.80)	118.88 (17.43)	<0.001
Diastolic blood pressure (mm Hg)	83.72 (12.48)	82.38 (11.15)	78.42 (10.14)	<0.001
FPG (mmol/L)	8.35 (3.15)	5.68 (0.61)	5.01 (0.51)	<0.001
Postprandial plasma glucose (mmol/dL)	14.93 (4.83)	7.85 (1.91)	5.98 (1.58)	<0.001
Cholesterol (mmol/L)	6.15 (1.38)	5.82 (1.29)	5.52 (1.14)	<0.001
HDL (mmol/L)	1.06 (0.30)	1.05 (0.26)	1.08 (0.28)	0.056
Triglycerides (mmol/L)	2.53 (1.86, 3.58)	2.12 (1.45, 3.04)	1.69 (1.16, 2.37)	<0.001

Abbreviations: BMI, Body mass index; WC, Waist circumference; FPG, Fasting plasma glucose; HDL, High density lipoprotein.
[a] P value was calculated using Student's t-test for continuous variables and the chi-square test for categorical variables and Man-Whitney U statistic for skewed variables.
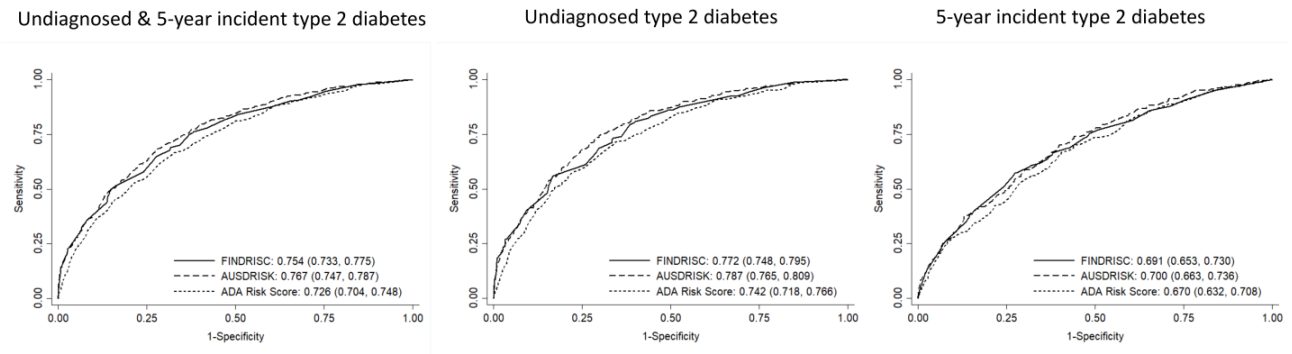
Figure 1. Receiver Operating Characteristic Curves.

< 0.001) models. Similarly, for identifying undiagnosed type 2 diabetes at baseline, AUSDRISK had the highest discrimination power (AUC = 0.79) as compared to FINDRISC (AUC = 0.77, *P* value = 0.033) and ADA (AUC = 0.74, *P* value < 0.001); similarly, AUSDRISK had higher discrimination power for identifying those with 5-year incident type 2 diabetes (AUC = 0.70) as compared with ADA risk prediction model (AUC = 0.67, *P* value = 0.037); however, there was no statistically significant difference between the discrimination powers of AUSDRISK (AUC = 0.70) and FINDRISC (AUC =

0.69, *P* value = 0.42) models.

Figure 2 shows the calibration plot of the original and calibrated models. The original model for AUSDRISK had a reasonable calibration for identifying those with undiagnosed type 2 diabetes or 5-year incident type 2 diabetes (chi-square = 17) while the original models for FINDRISC and ADA risk prediction models underestimated the risk with chi-square >1000. Recalibrated models for FINDRISC and ADA, however, had reasonably good calibration with chi-square <20. Similarly, calibration of the models for other outcomes
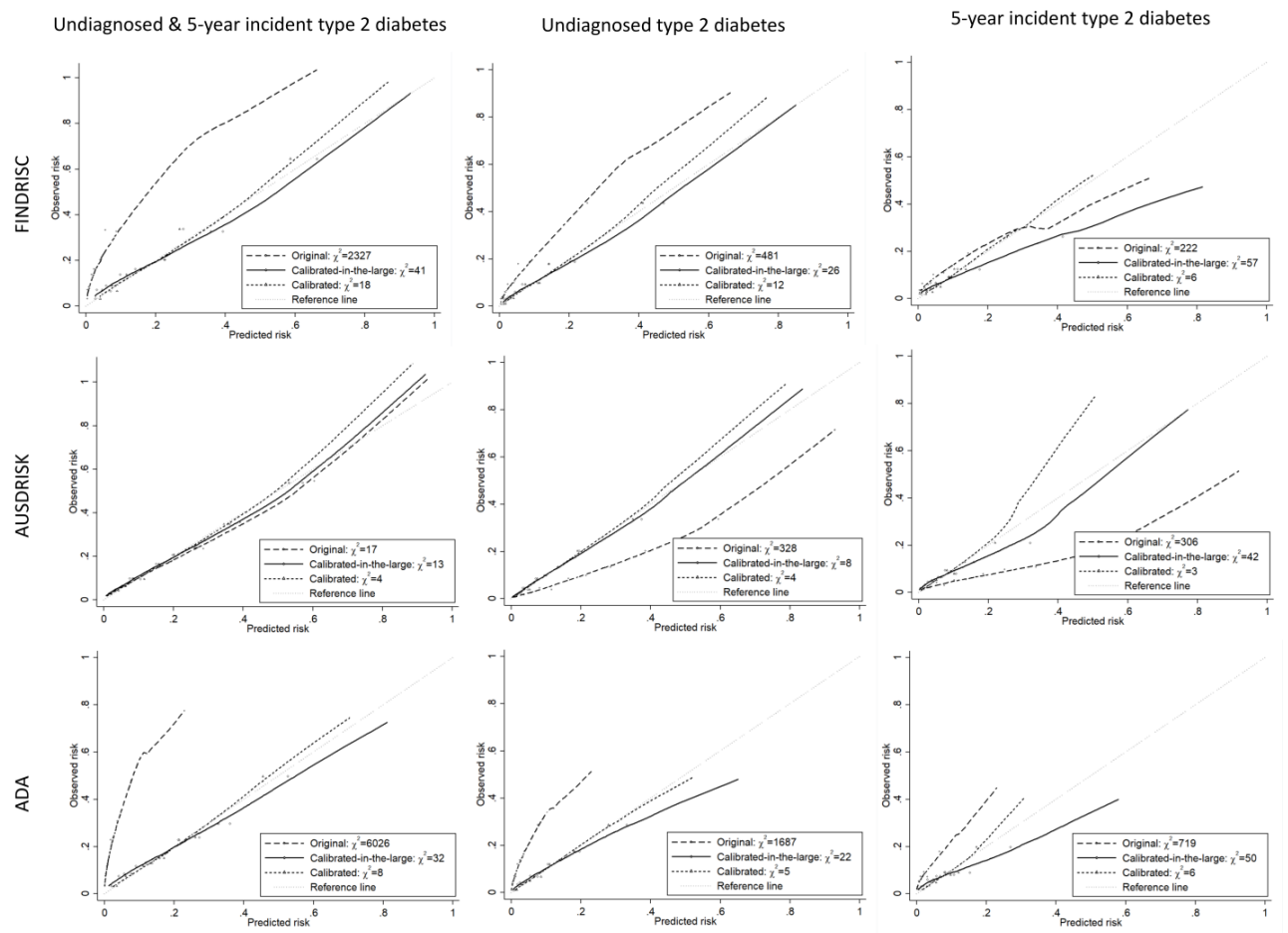

Figure 2. The Calibration Plot of the Risk Models*. Legend: *Markers show the observed vs. predicted values across the deciles of the predicted risks.

considerably reduced their chi-square below 20. Table S1 compares the equations for the original and calibrated models.

Figure 3 shows the decision curves of the models. Useful risk threshold ranges for the original models of FINDRISC and AUSDRISK in identifying those with undiagnosed type 2 diabetes or 5-year incident type 2 diabetes were 17%–42% and 9%–71%, respectively. The original model for ADA was not useful in any risk threshold. Calibration increased the usefulness of the models considerably with useful risk threshold ranges for the recalibrated models of FINDRISC, AUSDRISK, and ADA in identifying those with undiagnosed type 2 diabetes or 5-year incident type 2 diabetes being 10%–63%, 8%–64%, and 10%–42%, respectively. Similarly, the calibrated models were useful in a wide range of risk thresholds for identifying those with undiagnosed type 2 diabetes; however, they were useful in a narrow range of risk thresholds for identifying those with 5-year incident type 2 diabetes with useful risk threshold ranges for the recalibrated models of FINDRISC, AUSDRISK, and ADA in identifying those with 5-year incident type 2 diabetes being 6%–10%, 6%–9%, and 6%–8%, respectively.

The findings from sensitivity analyses regarding the discrimination (see Figure S1), calibration (see Figure S2)

and net benefit (see Figure S3) of the models were similar to those from the primary analyses.

Discussion

Summary of Findings

To the best of our knowledge, this is the first study to investigate the clinical utility of AUSDRISK, FINDRISC, and ADA risk prediction models for identifying individuals with undiagnosed type 2 diabetes or at high risk of developing it. Our findings showed that these three non-invasive diabetes risk prediction models are valid and reliable tools for identifying individuals with undiagnosed type 2 diabetes or at high risk of developing it. These three models had an acceptable discrimination ability, with AUSDRISK showing the highest discrimination ability for identifying those with undiagnosed type 2 diabetes or 5-year incident type 2 diabetes (all *P* values < 0.05). The original model for AUSRISK had an acceptable calibration for identifying those with undiagnosed type 2 diabetes or 5-year incident type 2 diabetes; FINDRISK and ADA risk prediction models reached acceptable calibration after recalibrating their intercept and slope. Calibrated models were useful in wide range of risk thresholds for identifying those with either undiagnosed type 2 diabetes or 5-year incident type 2 diabetes as well as identifying those with
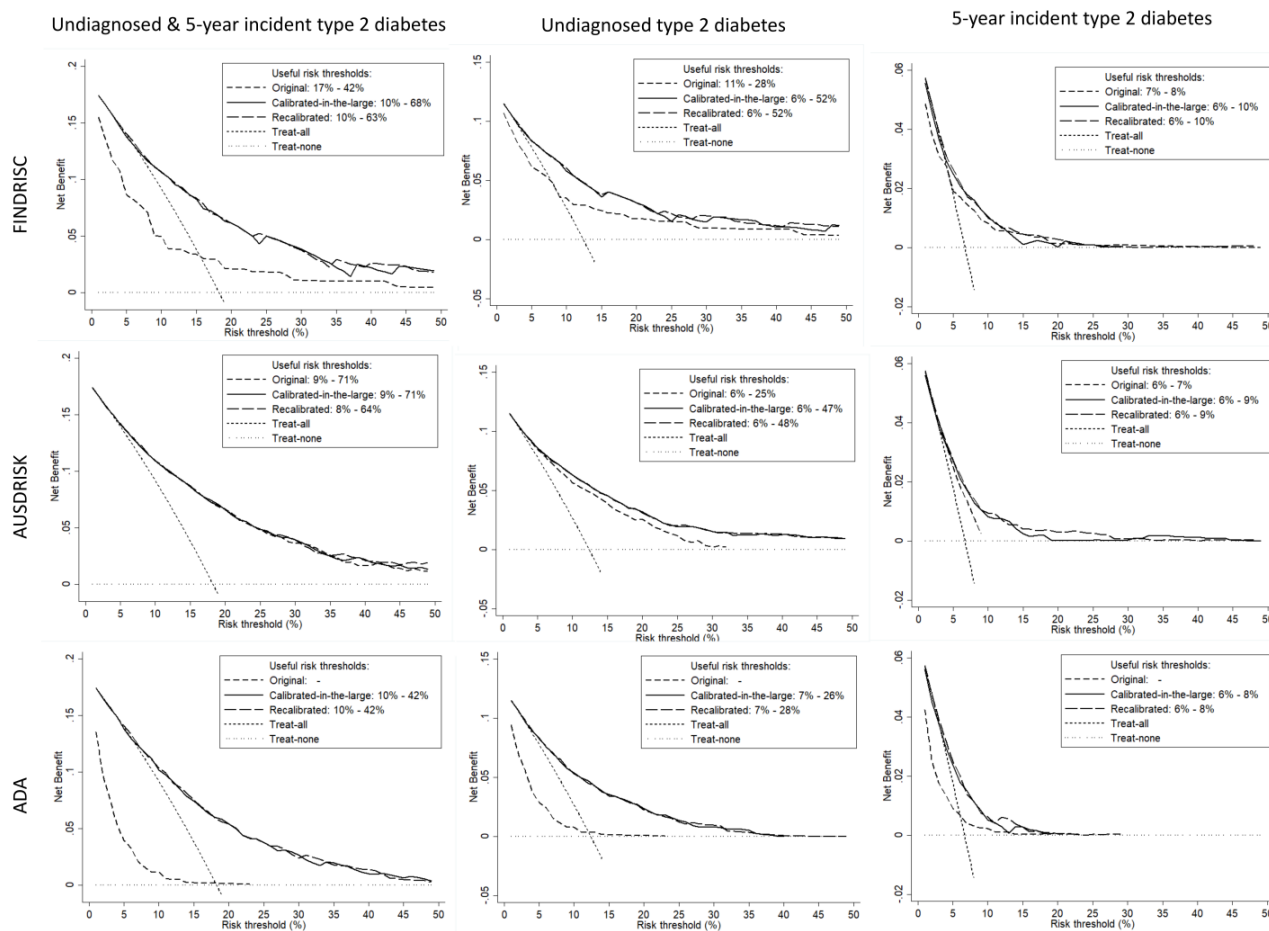

Figure 3. Decision Curves of Risk Models.

undiagnosed type 2 diabetes per se; however, ranges of useful risk thresholds for identifying those with incident type 2 diabetes were narrow per se.

Discrimination Power and Calibration of Risk Models

The discrimination ability of the AUSDRISK and ADA risk prediction models for the main outcome in this study is similar to the discrimination ability obtained in their original studies (0.77 vs. 0.78 for AUSDRISK and 0.73 vs. 0.74 for ADA risk prediction model); however, the discrimination ability for FINDRISC in this study is considerably lower as compared to the original study for FINDRISC (0.75 vs. 0.86).[16-18] Furthermore, the reported discrimination ability of the current study is comparable with other independent external validation studies showing the similar range of AUCs[14]; these findings support that the predictive value of the predictors in our population is similar to those populations that have been used to develop and validate the risk prediction models. Secondary analyses showed that the discrimination power of the models in identifying those with undiagnosed type 2 diabetes was considerably higher than their discrimination power in identifying those with 5-year incident type 2 diabetes.

Regarding the calibration of the models, the original model for AUSDRISK had an acceptable calibration for the main outcome of this study and original models for FINDRISC, and ADA risk prediction model achieved an acceptable calibration after calibrating the intercept and slope of the models. The models also achieved an acceptable calibration for detecting secondary outcomes after calibration for their intercept and slope. The need for calibration might be due to the higher prevalence of undiagnosed type 2 diabetes as well as its incidence rate in our population. In our population, 17.7% of our samples were found to have type 2 diabetes either at baseline or during the 5 years of follow-up, which is noticeably higher than that observed in the original studies of FINDRISC (4.1%), AUSDRISK (6.0%), and ADA (4.0%) risk prediction models.[16-18]

Clinical Utility of Risk Models

We showed the calibrated risk prediction models are useful in a wide range of the harm-to-benefit ratios for identifying those with undiagnosed type 2 diabetes or high risk individuals as well as for screening merely for undiagnosed cases; however, our findings showed that they are not useful for screening merely those at high risk of developing type 2 diabetes in most harm-to-benefit ratios. Based on our results, we recommend that policymakers consider using non-invasive risk prediction model in screening for those with undiagnosed type 2 diabetes or at high risk of developing it; these models are also useful in screening for undiagnosed type 2 diabetes per se. In case of screening for just those at high risk of type 2 diabetes, we

recommend further research to identify a useful screening method. Previous studies suggested that stepwise methods that combine a non-invasive risk prediction model with another invasive measurement can identify those at high risk of type 2 diabetes with high performance.[36-39]

This study does not recommend any specific cut-off for the predicted risk. It should be noted that a cut-off of risk for type 2 diabetes should be drawn considering the feasibility of different treatments and their efficacy as well as contextual factors; so it can vary considerably in different settings.[40] Due to variance in the prevalence of diseases, costs, and feasibility of treatments in different settings, previous literature suggested that the cut-off value for a diagnostic test is not universal and should be determined for each region and for each disease condition based on the harms and benefits of the screening and treatment.[40] It can be recommended to national policymakers in Iran to develop a clear cut-off of risk for screening for identifying individuals at high risk of type 2 diabetes as has been done by the Centers for Disease Control and Prevention in the United States.[41]

The better calibration and discrimination power in AUSDRISK might be explained by several factors. First, out of three risk prediction models used in this study, only AUSDRISK used ethnicity in its model.[17] The risk of diabetes was shown to be higher among Asians, Hispanics, and blacks compared to whites, before and after accounting for other type 2 diabetes risk factors.[17] AUSDRISK was the only risk prediction model that defined lower cut-offs for WC for Asian ethnicities.[17] Another reason for the better performance of AUSDRISK could be the higher number of variables used in its' model (n = 10) as compared to FINDRISC (n = 5), and ADA risk prediction model (n = 6).[16-18] Moreover, the outcome of this study was undiagnosed type 2 diabetes at baseline and 5-year incident type 2 diabetes based on the FPG, 2h-PG and treatment history which is very similar to the outcome of the original study of AUSDRISK,[17] while the original studies of FINDRISC and ADA had different follow-up periods and methods for defining their outcomes. It is important to notice that the aim of this study was not to replicate the findings of the original papers of risk prediction models but to assess whether these risk prediction model can be used in screening for type 2 diabetes in Iranian populations.

Strengths and Limitations

This study has some strengths and limitations. For strengths, to the best of our knowledge, it is the first study that investigated the clinical utility of three well-known risk prediction models for type 2 diabetes and it is one of the few studies in the low- and middle-income countries assessing the validation and calibration of these risk prediction models. We assessed the performance of the risk prediction models in an under-studied population

confronting the high and increasing burden of type 2 diabetes. The large sample size is one of the main strengths of this study, leading to high precision in our estimates. As for limitations, we had a drop-out rate of 25%. Moreover, for parental history of diabetes and physical activity, our definition differed for the original risk prediction model studies, e.g. instead of using the parental history of diabetes to derive AUSDRISK, we used family history of diabetes as data on the parental history of diabetes which had not been collected in the TLGS study. We assessed the performance of the models for identifying those with undiagnosed type 2 diabetes or 5-year incident type 2 diabetes based on FPG, 2h-PG, and treatment history among participants, aged ≥30 years, while the original risk prediction models have been developed for different outcomes and in samples with different age ranges. It is important to notice that the aim of this study was to identify a risk model that can address the current needs for a valid and reliable screening tool in an Iranian population and not to replicate the findings of the original studies of risk prediction models. We used Hosmer–Lemeshow chi-square to assess the calibration of the models, while some evidence showed that this test has a high rejection rate of acceptable models when large samples are used.[42]

In conclusion, we found that all of the three diabetes risk prediction models, AUSDRISK, FINDRISC, and ADA risk prediction modes had an acceptable performance for identifying individuals with undiagnosed type 2 diabetes or at high-risk of type 2 diabetes as well as only those with undiagnosed type 2 diabetes. We showed that they are useful in a wide range of harm-to-benefit ratios although they had limited utility in identifying those at high risk of developing type 2 diabetes in 5 years. Further research is needed to identify a useful screening method for prediction of type 2 diabetes. We recommend the original model of AUSDRISK for identifying individuals with undiagnosed type 2 diabetes or at high-risk of type 2 diabetes in the Iranian population because of its better performance as compared to other risk prediction models.

References

1. International Diabetes Federation 2017. IDF Diabetes Atlas. Online version IDF Diabetes Atlas. Available from: URL: www.diabetesatlas.org.
2. Harati H, Hadaegh F, Saadat N, Azizi F. Population-based incidence of Type 2 diabetes and its associated risk factors: results from a six-year cohort study in Iran. BMC Public Health. 2009;9:186. doi: 10.1186/1471-2458-9-186.
3. Sun Y, You W, Almeida F, Estabrooks P, Davy B. The effectiveness and cost of lifestyle interventions including nutrition education for diabetes prevention: a systematic review and meta-analysis. J Acad Nutr Diet. 2017;117(3):404-421.e36. doi: 10.1016/j.jand.2016.11.016.
4. Zhang X, Imperatore G, Thomas W, Cheng YJ, Lobelo F, Norris K, et al. Effect of lifestyle interventions on glucose regulation among adults without impaired glucose tolerance or diabetes: a systematic review and meta-analysis. Diabetes Res Clin Pract. 2017;123:149-164. doi: 10.1016/j.diabres.2016.11.020.
5. Li R, Qu S, Zhang P, Chattopadhyay S, Gregg EW, Albright A, et al. Economic Evaluation of Combined Diet and Physical Activity Promotion Programs to Prevent Type 2 Diabetes Among Persons at Increased Risk: A Systematic Review for the Community Preventive Services Task Force HHS Public Access. Ann Intern Med. 2015;163(6):452-60. doi:10.7326/M15-0469
6. Balk EM, Earley A, Raman G, Avendano EA, Pittas AG, Remington PL. Combined diet and physical activity promotion programs to prevent type 2 diabetes among persons at increased risk: a systematic review for the Community. Ann Intern Med. 2015;163(6):437-51. doi: 10.7326/M15-0452.
7. Haw JS, Galaviz KI, Straus AN, Kowalski AJ, Magee MJ, Weber MB, et al. Long-term sustainability of diabetes prevention approaches: A systematic review and meta-analysis of randomized clinical trials. JAMA Intern Med. 2017;177(12):1808-1817. doi: 10.1001/jamainternmed.2017.6040.
8. Diabetes Canada Clinical Practice Guidelines Expert Committee, Ekoe JM, Goldenberg R, Katz P. Screening for Diabetes in Adults. Can J diabetes. 2018;42 Suppl 1:S16-S19. doi:10.1016/j.jcjd.2017.10.004
9. Einarson TR, Bereza BG, Acs A, Jensen R. Systematic literature review of the health economic implications of early detection by screening populations at risk for type 2 diabetes. Curr Med Res Opin. 2017;33(2):331-358. doi: 10.1080/03007995.2016.1257977.
10. Alberti KGMM, Zimmet P, Shaw J. IDF Consensus on Type 2 Diabetes Prevention. Diabet Med. 2007;24(5):451-463.
11. Paulweber B, Valensi P, Lindström J, Lalic NM, Greaves CJ, McKee M, et al. A European evidence-based guideline for the prevention of type 2 diabetes. Horm Metab Res. 2010;42 Suppl 1:S3-36. doi: 10.1055/s-0029-1240928.
12. Putting Prevention First–NHS Health Check: Vascular Risk Assessment and Management Best Practice Guidance. 2009. Available from: https://www.theisn.org/images/

taskforce/United%20Kingdom/Putting%20Prevention%20First_NHS%20Health%20Check_Vascular%20Risk%20Assessments%20and%20Management%20Best%20Practice.pdf.

13. aaristo T, Peltonen M, Keinänen-Kiukaanniemi S, Vanhala M, Saltevo J, Niskanen L, et al. National type 2 diabetes prevention programme in Finland: FIN-D2D. Int J Circumpolar Health. 2007;66(2):101-12.

14. Kengne AP, Beulens JW, Peelen LM, Moons KG, van der Schouw YT, Schulze MB, et al. Non-invasive risk scores for prediction of type 2 diabetes (EPIC-InterAct): a validation of existing models. Lancet Diabetes Endocrinol. 2014;2(1):19-29. doi: 10.1016/S2213-8587(13)70103-7.

15. Abbasi A, Peelen LM, Corpeleijn E, van der Schouw YT, Stolk RP, Spijkerman AM, et al. Prediction models for risk of developing type 2 diabetes: Systematic literature search and independent external validation study. BMJ. 2012;345(7875):e5900. doi:10.1136/bmj.e5900

16. Lindström J, Tuomilehto J. The diabetes risk score. Diabetes Care. 2003;26(3):725-731.

17. Chen L, Magliano DJ, Balkau B, Colagiuri S, Zimmet PZ, Tonkin AM, et al. AUSDRISK: an Australian Type 2 Diabetes Risk Assessment Tool based on demographic, lifestyle and simple anthropometric measures. Med J Aust. 2010;192(4):197.

18. Bang H, Edwards AM, Bomback AS, Ballantyne CM, Brillon D, Callahan MA, et al. A patient self-assessment diabetes screening score: development, validation, and comparison to other diabetes risk assessment scores. Ann Intern Med. 2009 Dec 1;151(11):775-83. doi: 10.7326/0003-4819-151-11-200912010-00005.

19. Omech B, Mwita JC, Tshikuka JG, Tsima B, Nkomazna O, Amone-P'Olak K. Validity of the Finnish Diabetes Risk Score for detecting undiagnosed type 2 diabetes among general medical outpatients in Botswana. J Diabetes Res. 2016;2016:4968350. doi:1 0.1155/2016/4968350

20. Heltberg A, Andersen JS, Sandholdt H, Siersma V, Kragstrup J, Ellervik C. Predictors of undiagnosed prevalent type 2 diabetes–The Danish General Suburban Population Study. Prim Care Diabetes. 2018;12(1):13-22. doi: 10.1016/j.pcd.2017.08.005.

21. Dhippayom T, Chaiyakunapruk N, Krass I. How diabetes risk assessment tools are implemented in practice: A systematic review. Diabetes Res Clin Pract. 2014 ;104(3):329-42. doi: 10.1016/j.diabres.2014.01.008.

22. Krass I, Carter R, Mitchell B, Mohebbi M, Shih STF, Trinder P, et al. Pharmacy Diabetes Screening Trial: protocol for a pragmatic cluster-randomised controlled trial to compare three screening methods for undiagnosed type 2 diabetes in Australian community pharmacy. BMJ Open. 2017;7(12):e017725. doi: 10.1136/bmjopen-2017-017725.

23. Muchira JM, Stuart-Shor EM. Using Noninvasive Tools to Screen for Prediabetes and Undiagnosed Diabetes: Value Added to Cardiovascular Disease Prevention. J Cardiovasc Nurs. 2016;31(6):488-491. doi:1 0.1097/JCN.0000000000000372

24. Noble D, Mathur R, Dent T, Meads C, Greenhalgh T. Risk models and scores for type 2 diabetes: systematic review. BMJ. 2011;343:d7163. doi: 10.1136/bmj.d7163.

25. Janghorbani M, Adineh H, Amini M. Finnish Diabetes Risk Score to predict type 2 diabetes in the Isfahan diabetes prevention study. Diabetes Res Clin Pract. 2013;102(3):202-9. doi: 10.1016/j.diabres.2013.10.018.

26. Azizi F, Ghanbarian A, Momenan AA, Hadaegh F, Mirmiran P, Hedayati M, et al. Prevention of non-communicable disease in a population in nutrition transition: Tehran Lipid and Glucose Study phase II. Trials. 2009;10:5. doi:10.1186/1745-6215-10-5

27. Harati H, Hadaegh F, Momenan AA, Ghanei L, Bozorgmanesh MR, Ghanbarian A, et al. Reduction in incidence of type 2 diabetes by lifestyle intervention in a middle eastern community. Am J Prev Med. 2010;38(6):628-636.e1. doi: 10.1016/j.amepre.2010.03.003.

28. Bonita R. Surveillance of risk faktor for non-communicable disease; the WHO stepwise approach. Geneva: WHO: 2001. Available from: https://www.who.int/ncd_surveillance/media/en/269.pdf.

29. Ainsworth BE, Jacobs DR Jr, Leon AS. Validity and reliability of self-reported physical activity status: the Lipid Research Clinics questionnaire. Med Sci Sports Exerc. 1993;25(1):92-8.

30. Van der Heijden GJ, Donders AR, Stijnen T, Moons KG. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: a clinical example. J Clin Epidemiol. 2006;59(10):1102-9.

31. Steyerberg E. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Switzerland: Springer Science & Business Media; 2008.

32. D'Agostino RB Sr, Grundy S, Sullivan LM, Wilson P; CHD Risk Prediction Group. Validation of the Framingham Coronary Heart Disease Prediction Scores. JAMA. 2001;286(2):180-7. doi:10.1001/jama.286.2.180

33. Collins GS, Reitsma JB, Altman DG, Moons K. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. BMC Med. 2015;13(1):1. doi:10.1186/s12916-014-0241-z

34. Vickers AJ, Elkin EB. Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. Med Decis Making. 2006;26(6):565-74.

35. Baker SG, Van Calster B, Steyerberg EW. Evaluating a new marker for risk prediction using the test tradeoff: An update. Int J Biostat. 2012;8(1). pii: /j/ijb.2012.8.issue-1/1557-4679.1395/1557-4679.1395.xml. doi: 10.1515/1557-4679.1395.

36. Khunti K, Gillies CL, Taub NA, Mostafa SA, Hiles SL, Abrams KR, et al. A comparison of cost per case detected of screening strategies for Type 2 diabetes and impaired glucose regulation: modelling study. Diabetes Res Clin Pract. 2012;97(3):505-13. doi: 10.1016/j.diabres.2012.03.009.

37. Lee CMY, Versace VL, Malo JA, Shaw JE, Dunbar JA, Colagiuri S. Screening for diabetes prevention with diabetes risk scores–A balancing act. Diabetes Res Clin Pract. 2018;135:120-127. doi: 10.1016/j.diabres.2017.11.009.

38. Khunti K, Gillies CL, Taub NA, Mostafa SA, Hiles SL, Abrams KR, et al. A comparison of cost per case detected of screening strategies for type 2 diabetes and impaired glucose regulation: modelling study. Diabetes Res Clin Pract. 2012;97(3):505-13. doi: 10.1016/j.diabres.2012.03.009.

39. Hippisley-Cox J, Coupland C. Development and validation of QDiabetes-2018 risk prediction algorithm to estimate future risk of type 2 diabetes: cohort study. BMJ. 2017;359:j5019. doi: 10.1136/bmj.j5019.

40. Habibzadeh F, Habibzadeh P, Yadollahie M. On determining the most appropriate test cut-off value: the case of tests with continuous results. Biochem Med (Zagreb). 2016;26(3):297-307.

41. Albright AL, Gregg EW. Preventing type 2 diabetes in communities across the U.S.: The national diabetes prevention program. Am J Prev Med. 2013;44(4 Suppl 4):S346-S351. doi:10.1016/j.amepre.2012.12.009

42. Bartley A. Evaluating goodness-of-fit for a logistic regression model using the Hosmer-Lemeshow test on samples from a large data set. THESIS Submitt Partial Fulfillment Requir Degree Master Sci Grad Sch Ohio State Univ. 2014:1-30. doi:10.1007/s13398-014-0173-7.2