Research Paper

Non-lab and semi-lab algorithms for screening undiagnosed diabetes: A cross-sectional study

Wei Li [a,1], Bo Xie [a,1], Shanhu Qiu [a], Xin Huang [b], Juan Chen [a], Xinling Wang [c], Hong Li [d], Qingyun Chen [e], Qing Wang [f], Ping Tu [g], Lihui Zhang [h], Sunjie Yan [i], Kaili Li [j], Jimilanmu Maimaitiming [c], Xin Nian [d], Min Liang [e], Yan Wen [f], Jiang Liu [g], Mian Wang [h], Yongze Zhang [i], Li Ma [j], Hang Wu [a], Xuyi Wang [a], Xiaohang Wang [a], Jingbao Liu [a], Min Cai [a], Zhiyao wang [k], Lin Guo [k], Fangqun Chen [k], Bei Wang [b], Sandberg Monica [l], Per-Ola Carlsson [l,**], Zilin Sun [a,*]

[a] Department of Endocrinology, Zhongda Hospital, Institute of Diabetes, School of Medicine, Southeast Uiversity, Nanjing, China
[b] School of Public Health, Southeast University, Nanjing, China
[c] Department of Endocrinology, People's Hospital of Xinjiang Uyghur Autonomous Region, Urumqi, China
[d] Department of Endocrinology, First Affiliated Hospital of Kunming Medical University, Kunming, China
[e] Department of Endocrinology, The First Affiliated Hospital of Guangxi Medical University, Nanning, China
[f] Department of Endocrinology, China-Japan Union Hospital of Jilin University, Changchun, China
[g] Department of Endocrinology, The Third Hospital of Nanchang, Nanchang, China
[h] Department of Endocrinology, The Second Hospital of Hebei Medical University, Shijiazhuang, China
[i] Department of Endocrinology, The First Affiliated Hospital of Fujian Medical University, Diabetes Research Institute of Fujian Province, Fuzhou, China
[j] Department of Endocrinology, Xinjiang Uygur Autonomous Region Hospital of traditional Chinese Medicine, Urumqi, China
[k] Suzhou MetroHealth Medical Technology, Co., LTD, Suzhou, China
[l] Department of Medical Cell Biology, Uppsala University, SE-75123 Uppsala, Sweden

A R T I C L E   I N F O

A B S T R A C T

Background: The terrifying undiagnosed rate and high prevalence of diabetes have become a public emergency. A high efficiency and cost-effective early recognition method is urgently needed. We aimed to generate innovative, user-friendly nomograms that can be applied for diabetes screening in different ethnic groups in China using the non-lab or noninvasive semi-lab data.

Methods: This multicenter, multi-ethnic, population-based, cross-sectional study was conducted in eight sites in China by enrolling subjects aged 20–70. Sociodemographic and anthropometric characteristics were collected. Blood and urine samples were obtained 2 h following a standard 75 g glucose solution. In the final analysis, 10,794 participants were included and randomized into model development (n = 8096) and model validation (n = 2698) group with a ratio of 3:1. Nomograms were developed by the stepwise binary logistic regression. The nomograms were validated internally by a bootstrap sampling method in the model development set and externally in the model validation set. The area under the receiver operating characteristic curve (AUC) was used to assess the screening performance of the nomograms. Decision curve analysis was applied to calculate the net benefit of the screening model.

Results: The overall prevalence of undiagnosed diabetes was 9.8% (1059/10794) according to ADA criteria. The non-lab model revealed that gender, age, body mass index, waist circumference, hypertension, ethnicities, vegetable daily consumption and family history of diabetes were independent risk factors for diabetes. By adding 2 h post meal glycosuria qualitative to the non-lab model, the semi-lab model showed an improved Akaike information criterion (AIC: 4506 to 3580). The AUC of the semi-lab model was statistically larger than the non-lab model (0.868 vs 0.763, P < 0.001). The optimal cutoff probability in semi-lab and non-lab nomograms were 0.088 and 0.098, respectively. The sensitivity and specificity were 76.3% and 81.6%, respectively in semi-lab nomogram, and 72.1% and 67.3% in non-lab nomogram at the optimal cut off point. The decision curve analysis also revealed a bigger decrease of avoidable OGTT test (52 per 100 subjects) in the semi-lab model compared to the non-lab model (36 per 100 subjects) and the existed New Chinese Diabetes Risk Score (NCDRS, 35 per 100 subjects).

Conclusion: The non-lab and semi-lab nomograms appear to be reliable tools for diabetes screening, especially in

---

* Correspondence to: Z.L. Sun, Department of Endocrinology, Zhongda Hospital, Institute of Diabetes, School of Medicine, Southeast University, Nanjing, China.
** Correspondence to: P. O. Carlsson, Department of Medical Cell Biology, Uppsala University, Uppsala, Sweden.
E-mail addresses: per-ola.carlsson@mcb.uu.se (P.-O. Carlsson), sunzilin1963@126.com (Z. Sun).
[1] These two authors contributed equally to this paper.

developing countries. However, the semi-lab model outperformed the non-lab model and NCDRS prediction systems and might be worth being adopted as decision support in diabetes screening in China.

Research in context

Evidence before this study

We searched PubMed with the terms "diabetes", "risk score", "nomogram", "ethnic groups", "urine glucose", "decision curve analysis", "validation", "net benefit" and "risk algorithm" in English before December 30, 2017. Although several risk algorithms for diabetes had already been developed by the researchers all over the world. Most of which were not suitable for Chinese population. Although Zhou and colleagues developed a risk algorithm for Chinese population based on the China National Diabetes and Metabolic Disorders Study, ethnographic heterogeneity was not take into consideration in this risk score. Besides, none of the existing risk scores for diabetes was evaluated with clinical usefulness and the economic benefit. Before the preparation of this research, a study using postprandial 2-h period urine glucose to screen for diabetes in Chinese natural population was performed among 7485 subjects in six cities of Jiangsu. Postprandial 2-h period urine glucose showed a good discriminability for diabetes screening.

Added value of this study

We developed two simple-to-use nomograms (Non-lab and Semi-lab nomogram) and two accordingly websites that use data commonly available in clinical practice to estimate risk of diabetes in different ethnic groups. We also evaluated the clinical usefulness of the two risk algorithms by the decision curve analysis. Our non-lab model reveled similar discriminability when compared with the existed risk algorithm (The New Chinese Diabetes Risk Score). Our semi-lab model showed the best performance in risk prediction and clinical benefit among all the three models in this report.

Implications of all the available evidence

Urine glucose alone may not suitable for diabetes screening, But the high specificity of urine glucose may compromise the poor efficiency of the risk model derived from descriptive clinical data. Our study verified this hypothesis. With the advantages of non-invasive, cheap postprandial 2-h period glycosuria qualitative, our semi-lab model may be suitable for large scale screening of diabetes and also self-assessment of diabetes risk.

1. Introduction

Diabetes becomes one of the largest health emergencies worldwide. The number of people with this life-changing condition is increasing dramatically year by year. It is estimated that one in 11 adults aged 20–79 have diabetes worldwide, among whom nearly 50% [1] are left undiagnosed. This situation is even more terrifying in China. Nationally representative surveys among adults in mainland China revealed that the overall prevalence of diabetes in China ranged from 9.7% to 11.6% and among all patients with diabetes 60.7–69.9% individuals are unaware of their condition [2–4]. Moreover, evidence suggests that up to 25% of people with diabetes have developed microvascular complications at the time of diagnosis [5]. Thus, early detection of diabetes seems to be extremely important for this proportion of patients.

Fasting plasma glucose (FPG), 2-h postprandial glucose (2-h PG) after 75-g oral glucose tolerance tests (OGTT), or HbA1c level is recommended for diabetes screening [6]. But their practicability has been challenged by low efficacy [7], time-consuming, and high cost. Furthermore, these tests are invasive, which might compromise the compliance or willingness of individuals to screen for diabetes. Therefore, a simple and user-friendly assessment tool with comparable or superior sensitivity and specificity is urgently in need.

There are several mathematical models designed to identify diabetes [8–12], with the Finnish diabetes risk score [11] as the most feasible one. However, its high efficacy might be largely attributed to the question: "Have you ever been told by a health-care professional that you have diabetes or latent diabetes? No/Latent diabetes/Diabetes". Moreover, this tool might be not suitable for Chinese population because most people in China may not have periodic health visits, let alone being informed of diabetes. Although recently a New Chinese Diabetes Risk Score [8] was developed by Zhou and colleagues in 2013 for detecting diabetes in Chinese population, their algorithm was developed and validated in low-income residents, and its accuracy remains unclear in Chinese with different ethnic groups. Furthermore, no nomograms [13] which can pictorially depict an individual probability of undiagnosed diabetes have been developed in the mainland of China and no proper methods for the assessment of the clinical utility for a risk model have been reported.

The aims of this study was to generate innovative and user-friendly nomograms for the screening of diabetes in different ethnic groups of non-diabetic subjects in China by using the non-lab or semi-lab data and to assess these algorithms by decision curve analysis [14, 15] for their clinical utility at different threshold probabilities. In addition, this study was also aimed to compare the performance of non-lab and semi-lab model developed in this study with the New Chinese Diabetes Risk Score derived by Zhou in different ethnic groups to evaluate its effectiveness.

2. Material and methods

2.1. Data for development of algorithms

This observational Study on Evaluation of iNnovated Screening tools and determInation of optimal diagnostic cut-off points for type 2 diaBetes in Chinese muLti-Ethnic (SENSIBLE study) was conducted in 8 centers including six ethnic groups in 7 provinces in China from November 2016 to June 2017. A multi-stage cluster and simple randomization method was applied to recruit subjects aged 20–70 years in 1–2 provinces that were randomly selected from different regions (north, south, east, west, and central) across China, where Jilin and Hebei provinces (north China), Yunnan and Guangxi provinces (south China), Fujian provinces (east China), Xinjiang Uyghur autonomous region (west China), and Jiangxi provinces (central region) were finally chosen. Afterwards, city names in each province were numbered, which were chosen by simple random sampling method. Ten cities were finally enrolled in the first stage of sampling. In the second stage, 10 neighborhood communities and 10 administrative villages were randomly selected. Finally, individuals who lived at least 5 years in their

current residence were randomly sampled with stratifications on sex and age distributions. However, the following featured individuals were excluded: 1) refused to sign the informed consent; 2) pregnancy; 3) mental illness; 4) other physiological diseases unable to finish the procedures of this survey. A total of 13,620 subjects were invited and 12,017 subjects participated in this study, which give a response rate of 88.2%. After excluding 649 participants with self-reported diabetes, 538 with missing data on sociodemographic information (e.g. age, gender, ethnicities, family history of diabetes), physical examination characteristics (including waist circumference, height, weight, systolic blood pressure, diastolic blood pressure), or laboratory indices (including hemoglobin A1c, fasting plasma glucose, 2-h plasma glucose, postprandial 2-h period glycosuria qualitative, triglyceride, total cholesterol, high-density lipoprotein, low-density lipoprotein), and 36 outliers with waist circumference > 99.9 percentile(120 cm) or < 0.1 percentile(54 cm) or with BMI >99.9 percentile(41.42 kg/m^2) or < 0.1 percentile(15.60 kg/m^2), a total of 10,794 participants(Fig. 1) were included in the final data analysis.

This study protocol was approved by the Ethical Review Committees of Zhongda Hospital, Southeast University, and other participating institutes. Written informed consent was obtained from each participant before participation.

2.2. Procedures

The SENSIBLE study was conducted in each neighborhood community or administrative village with the help from co-operating grade IIIA hospital. All the research staff was trained by an experienced executive director in order to guarantee a unified standard procedure. All eligible participants were informed to maintain their usual lifestyle for at least 3 days and were fasted at least 10 h before blood sample withdrawn, which was used to measure HbA1c, FPG, postprandial 2-h period glycosuria qualitative, triglyceride (TG), total cholesterol (TC), high-density lipoprotein (HDL), low-density lipoprotein (LDL). Then these participants were instructed to empty their bladder and swallowed a standard 75 g glucose solution for an OGTT, with the
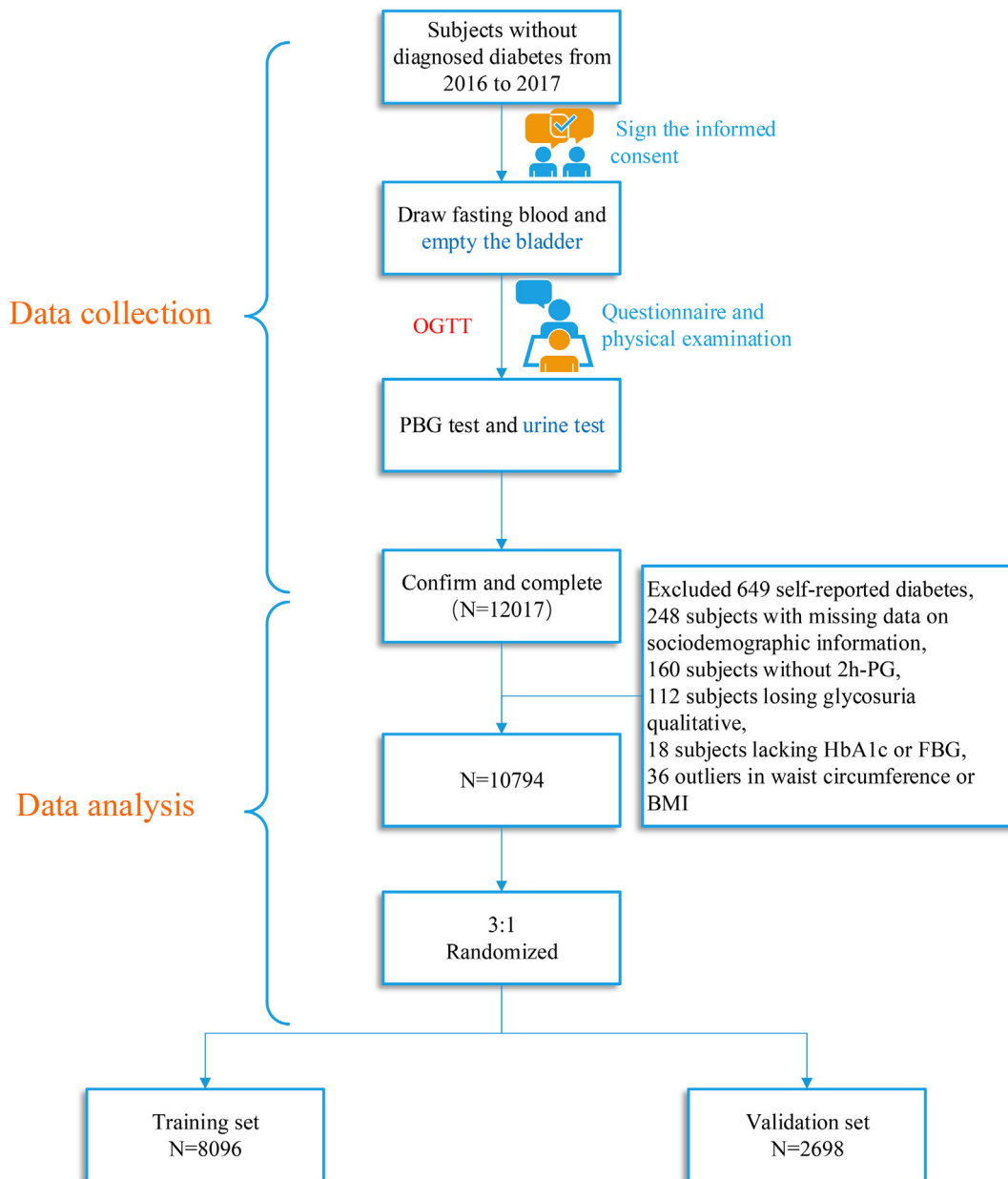

Fig. 1. Flow chart of the research.

blood sample taken 2 h later for 2 h-PG measurement. Immediately after the 2 h blood sample taken, participants were asked to empty their bladder again, collecting the 2 h period urine for glycosuria qualitative measurement which was conducted by an automatic urine analyzer (Uritest-500B, URIT Corporation, China). The glycosuria results were categorized as −, ±(trace), +, which represent the increasing concentration of urine glucose.

A structured questionnaire administered by trained interviewers was conducted to get information on sociodemographic characteristics, lifestyle factors, and medical history. Body weight, height, and waist circumference were measured with standardized protocols. Body mass index (BMI) was calculated as body weight (kg) divided by squared height (m). Blood pressure was measured 3 times at the non-dominant arm after 5 min of rest at a seated position using an automated device (YE680E, yuwell, China). All blood samples were centrifuged on site within 30 min after collection. For the serum and the whole blood samples, they were shipped at 4 °C by air to the central laboratory in Nanjing Adicon Clinical Laboratories. All the blood specimens were measured immediately after arrival. FPG, 2 h-PG, TG, TC, HDL, LDL were measured using an automatic chemistry analyzer (Synchron LX-20, Beckman Coulter Inc., CA, USA). HbA1c was measured with high-performance liquid chromatography (HPLC; D-10™ Hemoglobin Analyzer, Bio-Rad Inc., CA, USA).

2.3. Definitions

Diabetes was defined according to ADA 2015 criteria as 1) FPG ≥ 7.0 mmol/L, 2) 2 h-PG ≥ 11.1 mmol/L, 3) HbA1c concentration at a level of 6.5% or more.

2.4. Development and assessment of recognition models

The finally enrolled 10,794 participants were divided into a training set (N = 8096) and a validation set (N = 2698) at a ratio of 3:1 using a simple random sampling method. Data in the training set were used to develop the recognition models for undiagnosed diabetes. A multivariable binary logistic regression model using the backward stepwise method was applied to develop the algorithm. The dependent variable in the model was the undiagnosed diabetes, while the independent variables including: (1) sociodemographic information such as gender, age, ethnicities, family history of diabetes, education level, tea habit, sleep quality, sleep time, diet habit, exercise, vegetable daily consumption, and income; (2) physical examination characteristics such as waist circumference, BMI, uncontrolled blood pressure (SBP ≥ 140 mmHg and or DBP ≥ 90 mmHg). The stepwise process was evaluated by the Akaike information criterion (AIC) and Bayesian information criterion (BIC) statistics and only the model with the lowest AIC and BIC achieved was considered as the final one. Factors such as gender, age, ethnicities, BMI, waist circumference, hypertension, family history of diabetes, and vegetable daily consumption were included in the final non-lab model. The postprandial 2-h period glycosuria qualitative results were added to the non-lab model and the interaction between variables were also considered for the construction of the semi-lab model. Nomograms were constructed using the rms package in R software version 3.4.1 (http://www.r-project.org) according to the methods reported previously [16–19].

To achieve an unbiased estimate of our models, a bootstrap sampling method was used in the training set to validate the performance of our non-lab and semi-lab models internally. Then external validation was performed in the validation set (N = 2698) using the area under the receiver operating characteristic curve (AUC) and the differences between different AUCs were compared using the DeLong method [20]. Youden's index was used to find the optimal cut-off value for the detection of undiagnosed diabetes. The accuracy of our models was further verified and compared with the previously published New Chinese Diabetes Risk Score in different gender and ethnic groups using the ROC curve and

AUC. The clinical usefulness was evaluated using the net benefit (The average profit of a prediction which derived from the true-positive rate multiply the gain of doing a test or treatment minus the false-positive rate multiply the loss of doing a test or treatment) calculate by the following formula which was derived by

$$\text{Net benefit} = \frac{true-positive\ count}{n} - \frac{false-positive\ count}{n}\left(\frac{p_t}{1-p_t}\right)$$

Andrew J. Vickers [14], where n means the sample size, p_t means the threshold probability which can categorize the model derived probability into positive or negative, in this formula the gain is represented as 1 and the loss is represented as $p_t/1- p_t$. The decision curves of the non-lab and semi-lab model were plotted by the rmda package included in R software version 3.4.1 (http://www.r-project.org) according to the previously published papers [21]. Finally, two simple and user-friendly websites according to non-lab model (https://yunxuan.shinyapps.io/nonlabmodel/) and semi-lab model (https://yunxuan.shinyapps.io/semilabmodel/) were developed to estimate the individualized risk of diabetes.

2.5. Statistical analyses

An estimated 260 subjects would be needed to provide 99% power for a receiver operating characteristic study, assuming the area under the ROC is 0.75 and the prevalence of undiagnosed diabetes is 10%, with a two-sided α of 0.05.

All the statistical analysis was performed using the Empower Stats (www.empowerstats.com, X&Y solutions, Inc. Boston MA) and R software version 3.4.1 (http://www.r-project.org). The relevant packages including rms, rmda, pROC, DynNom and shiny. Continuous variables were described as means ± SD or median (25th–75th percentile) and categorical data were presented as number and percentage. The difference between the model development group, model validation group and total samples was compared using one-way analysis of variance (ANOVA) for continuous data and Chi-squared tests for categorical variables. Kruskal-Wallis test was applied for the variables with a skewed distribution.

2.6. Role of the funding source

The sponsors had no role in the study design; collection, data analysis, data interpretation or writing of the report. The corresponding author had full access to all the data and the final responsibility for the decision to submit for publication.

3. Results

There were 1059 cases with newly diagnosed diabetes based on ADA criteria (Table 1), accounting for 9.8% of the total population. Notably, the prevalence of undiagnosed diabetes differed across populations with different ethnicities. For example, the Han population had the highest prevalence of newly diagnosed diabetes, which was 13.7%, while the Kazak population had the lowest prevalence of 4.2%. The Zhuang (11.5%) and the Korean populations (11.5%) had a similar prevalence. The Uyghur adults and the Dai participants had a relatively low prevalence rate of 5.2% and 7.3%. There were no significant differences in sociodemographic characteristics, physical examination characteristics or laboratory characteristics between the model development and model validation groups (Table 1).

The best non-lab model generated from the backwards stepwise regression showed that gender (percentage of males), increased age, BMI, waist circumference, hypertension, family history of diabetes and low daily consumption of vegetables were potential risk factors for newly diagnosed diabetes (Table 2). Among them, family history was the strongest risk factor with an OR of 1.72 (95%CI 1.41–2.08). But

Table 1
Baseline characteristics of participants in different groups.

	Total (n = 10,794)	Training set (n = 8096)	Validation set (n = 2698)	P-Value
Age (years)	49.2 ± 12.5	49.1 ± 12.5	49.2 ± 12.4	0.960
Gender				0.970
Female	7398 (68.5%)	5554 (68.6%)	1844 (68.3%)	
Male	3396 (31.4%)	2542 (31.4%)	854 (31.7%)	
Ethnic Groups				0.967
Korean	1351 (12.5%)	999 (12.3%)	352 (13.0%)	
Dai	1949 (18.1%)	1442 (17.8%)	507 (18.8%)	
Han	3084 (28.6%)	2324 (28.7%)	760 (28.2%)	
Kazak	853 (7.9%)	635 (7.8%)	218 (8.1%)	
Uyghur	1682 (15.6%)	1271 (15.7%)	411 (15.2%)	
Zhuang	1875 (17.4%)	1425 (17.6%)	450 (16.7%)	
Vegetable daily consumption				0.990
Very low	39 (0.4%)	27 (0.3%)	12 (0.4%)	
Low	555 (5.1%)	415 (5.1%)	140 (5.2%)	
Normal	6975 (64.6%)	5242 (64.7%)	1733 (64.2%)	
High	3225 (29.9%)	2412 (29.8%)	813 (30.1%)	
Undiagnosed diabetes	1059 (9.8%)	779 (9.6%)	280 (10.4%)	0.520
Hypertension	3867 (35.8%)	2897 (35.8%)	970 (36.0%)	0.987
Family history of diabetes	1734 (16.1%)	1308 (16.2%)	426 (15.8%)	0.904
BMI(kg/m2)	24.8 ± 3.9	24.8 ± 3.9	24.9 ± 3.9	0.636
Waist circumference(cm)	82.6 ± 10.9	82.6 ± 10.9	82.7 ± 11.0	0.898
HbA1c(%)	5.5 ± 0.8	5.5 ± 0.8	5.5 ± 0.7	0.831
FPG(mmol/L)	5.5 ± 1.2	5.5 ± 1.2	5.5 ± 1.1	0.996
2 h-PG(mmol/L)	7.0 ± 3.1	7.0 ± 3.1	7.0 ± 3.2	0.979
Glycosuria qualitative				0.678
−	9445 (87.5%)	7097 (87.7%)	2348 (87.0%)	
+−	232 (2.1%)	171 (2.1%)	61 (2.3%)	
+	1117 (10.4%)	828 (10.2%)	289 (10.7%)	
TG(mmol/L)	1.2(0.8,1.8)	1.2(0.8,1.80)	1.2(0.8,1.80)	0.727
TC(mmol/L)	5.2 ± 1.1	5.2 ± 1.2	5.2 ± 1.1	0.984
LDL(mmol/L)	3.0 ± 0.9	3.0 ± 0.9	3.0 ± 0.8	0.980
HDL(mmol/L)	1.6 ± 0.4	1.6 ± 0.4	1.6 ± 0.4	0.948

Data are presented as n, n(%), mean ± SD or median(IQR).

considering the span of the factor (OR*[max(factor)-min(factor)]), age was dominated as the strongest factor in the non-lab model (Fig. 2). After adding postprandial 2-h period glycosuria qualitative to the non-lab model to generate the semi-lab model, it is interesting that male gender became a protective factor for newly diagnosed diabetes (OR 0.75, 95% CI 0.61–0.93). A subsequent logistic regression model involving the interaction of glycosuria and gender showed that there exists an interaction between these two variables (Table 2, Glycosuria + interactive with male gender, OR 0.61 and 95% CI 0.41–0.90, P = 0.01). The non-lab model presented a c-index of 0.763 (95% CI 0.747–0.780), while the final semi-lab model exhibited a bigger one (c-index 0.868, 95% CI 0.854–0.882).

The internal bootstrap validation demonstrated that at a probability between 0 and 0.25, the non-lab nomogram derived curve fitted well with the bias-corrected curve and the ideal curve. But when the probability was set to be higher than 0.25, the non-lab model may overestimate the probability of undiagnosed diabetes (Fig. 3a). Our semi-lab model resembled this tendency but the start point of overestimation at the predicted probability was higher than the non-lab model (Fig. 3b). Both our non-lab and semi-lab model showed a good fitting and calibration, with the mean absolute error being 0.004 for the non-lab model and 0.003 for the semi-lab model.

The non-lab and semi-lab nomograms were further validated using ROC internally in the training set and externally in the validation set. The AUC for the non-lab nomogram (Fig. 3c) in the training set was 0.763 (95% CI 0.747–0.780), yielding a sensitivity of 72.1% and a specificity of 67.3% at the optimal cutoff value (P = 0.098, P means model derived probability) that maximized the Youden's index. Yet in the validation set the AUC was 0.753 (95% CI: 0.726–0.781), along with a sensitivity of 84.3% and a specificity of 53.7% at the corresponding threshold. However, the AUCs for the semi-lab nomogram (Fig. 3d) in the training (0.868, 95% CI: 0.854–0.882) and validation sets (0.872, 95% CI: 0.848–0.897) were larger than those for the non-lab nomogram

(AUC = 0.763, and 0.753, respectively). At the optimal corresponding cutoff values (P = 0.088, P means model derived probability), the no-mogram yielded a sensitivity of 76.3% and a specificity of 81.6% for the training set and a sensitivity of 70.7% and a specificity of 90.1% for the validation set. Our semi-lab model showed a 14.3%(81.6%–67.3%) increment in specificity without lowering the sensitivity at the optimal cut off point in the training set.

The performances of the non-lab and semi-lab nomograms were compared with the New Chinese Diabetes Risk Score [8] in the whole populations (n = 10,794) as well as in those with different gender and ethnicities. As shown in Fig. 4. Our non-lab nomogram demonstrated similar accuracy, discriminability with the New Chinese Diabetes Risk Score no matter in the subgroup of gender or ethnic groups. Both the non-lab nomogram and the New Chinese Diabetes Risk Score were inferior to the semi-lab nomogram in accuracy and discrimination.

The decision curve analysis comparing the clinical usefulness of the non-lab and semi-lab models was shown in Fig. 5. The threshold probability for diabetes was plotted in x axis and the standard net benefit using the model was plotted in y axis. The area among the model curve, treat all line and treat none line represent the clinical usefulness of each model. In this analysis, the New Chinese Diabetes Risk Score, non-lab model, and semi-lab model all showed a better cost effective than treat all and treat none, and the semi-lab model exhibited the best performance. For example, at a threshold of 10%, the non-lab model and the New Chinese Diabetes Risk Score would cause a reduction of 36 and 35 subjects per 100 participants from performing OGTT while a big increase of 52 subjects per 100 participants (calculated [14] by (net benefit of the model – net benefit of treat all)/(p_t/(1 – p_t)) × 100) for semi-lab model, without increasing the number of false-positive results. This means our semi-lab model will save 16–17 subjects per 100 people from the cost of OGTT and also the human resources during the process of the test.

Table 2
Odds ratio (95% CI) and β-coefficient in non-lab model and semi-lab model estimated by logistic regression analysis using the data from the training set.

Factors	Non-lab(n = 8096)			Semi-lab(n = 8096)		
	β-Coefficient	P-Value	OR(95%CI)	β-Coefficient	P-Value	OR(95%CI)
Age(years)	0.05	<0.01	1.05(1.04–1.06)	0.06	<0.01	1.06(1.05–1.07)
Gender						
Female	—	—	1.00	—	—	1.00
Male	0.28	<0.01	1.32(1.11–1.57)	−0.07	0.01	0.94(0.72–1.21)
Ethnic groups						
Han	—	—	1.00	—	—	1.00
Korean	0.22	0.08	1.24(0.97–1.59)	0.15	0.31	1.16(0.87–1.53)
Dai	−0.01	0.93	0.99(0.76–1.28)	0.17	0.32	1.19(0.86–1.56)
Kazak	−1.42	<0.01	0.24(0.15–0.37)	−1.44	<0.01	0.24(0.14–0.38)
Uyghur	−0.83	<0.01	0.48(0.31–0.60)	−0.54	<0.01	0.58(0.40–0.82)
Zhuang	−0.03	0.78	0.97(0.78–1.20)	−0.06	0.64	0.94(0.73–1.21)
Vegetable daily consumption						
Very low	—	—	1.00	—	—	1.00
Low	−0.60	0.33	0.55(0.18–2.11)	−0.98	0.15	0.37(0.11–1.68)
Normal	−1.01	0.09	0.36(0.12–1.36)	−1.39	0.04	0.25(0.07–1.08)
High	−0.98	0.11	0.37(0.12–1.41)	−1.27	0.06	0.28(0.08–1.23)
Hypertension						
No	—	—	1.00	—	—	1.00
Yes	0.47	<0.01	1.60(1.36–1.90)	0.28	<0.01	1.32(1.09–1.60)
Family history of DM						
No	—	—	1.00	—	—	1.00
Yes	0.54	<0.01	1.72(1.41–2.08)	0.37	<0.01	1.45(1.16–1.81)
BMI(kg/m2)	0.08	<0.01	1.08(1.04–1.12)	0.09	<0.01	1.10(1.06–1.14)
Waist circumference(cm)	0.03	<0.01	1.03(1.02–1.04)	0.02	0.01	1.02(1.01–1.04)
Glycosuria qualitative						
−				—	—	1.00
+/−				1.62	<0.01	5.05(2.80–8.75)
+				2.28	<0.01	24.96(19.15–32.63)
Interactive effect						
Others				—	—	1.00
+/− *Gender = male				−0.19	0.65	0.82(0.36–1.89)
+ *Gender = male				−0.50	0.01	0.61(0.41–0.90)

AIC, Akaike information criterion; BIC, Bayesian information criterion; OR, odds ratio; BMI, body mass index. +/−*Gender = male means glycosuria qualitative +/− interactive with male gender, +*Gender = male means glycosuria qualitative + interactive with male gender.

4. Discussion

This SENSIBLE study explored the rate of undiagnosed diabetes in six different ethnic groups in China. According to this multicenter, multinational data, two models were developed for predicting undiagnosed diabetes. The non-lab model based on eight variables including gender, age, ethnicities, BMI, waist circumference, uncontrolled blood pressure, family history of diabetes, and vegetable daily consumption. Postprandial 2-h period glycosuria qualitative and interaction of gender with glycosuria was added to the non-lab model for the construction of the semi-lab model. Both models showed good discriminability in training set and validation set subjects. Comparisons of the two models with the New Chinese Diabetes Risk Score in the subgroup of gender or ethnicities suggested that the semi-lab model exhibited the best performance, while the non-lab model had a similar performance as the New Chinese Diabetes Risk Score. The semi-lab model also provided better clinical usefulness than the other two models as indicated by the decision curve analyses. Besides, this comparison further proved the practicability of the New Chinese Diabetes Risk Score in China. Of note, our Non-lab model and the existed New Chinese Diabetes Risk Score might be more suitable for individuals lacking medical resources, or for subjects with urinary tract infection or taking drugs that may affect the accuracy of glycosuria qualitative measurement, while our semi-lab model seems to be more appropriate for individuals under the guidance of community doctors or general practitioner and our semi-lab model is also considerable for the epidemiologic study on diabetes screening.

It is reported that there might exist a racial difference for assessing diabetes risk according to the multiethnic cohort study enrolling 59,824 nondiabetic adults from South Asian, Chinese, African, and Caucasian [22]. Partly in support of this and consistent with the published data [23, 24], our study indicated that the rate of undiagnosed diabetes was highest in Han population and lowest in Kazak people. Besides, this rate was investigated in another two extra nationalities of Zhuang and Dai which was not reported previously. However, there was no significant improvement comparing our ethnic-based non-lab algorithms with the New Chinese Diabetes Risk Score in different subgroups. This may be probably attributable to the differences in BMI and waist circumference among the six ethnic groups. For example, both the Kazak (undiagnosed diabetes rate 4.2%) and Uyghur population (undiagnosed diabetes rate 5.2%) were top two ethnic groups with high BMI (27.6 ± 5.0 kg/m^2 in Kazak and 25.6 ± 4.6 kg/m^2 in Uyghur) and waist circumference (90.1 ± 12.2 cm in Kazak and 88.5 ± 11.3 cm in Uyghur) while the Zhuang population (undiagnosed diabetes rate 11.5%) ranked last in BMI (23.7 ± 3.7 kg/m^2) and the Dai people (undiagnosed diabetes rate 7.3%) in waist circumference (76.3 ± 8.5 cm). These indicate that the multicollinearity may be occurred when the variables of ethnicities, BMI and waist circumference were included in the model. This multicollinearity may be responsible for the inaction of the non-lab model when applied in the subgroup of nationalities.

Although a number of diabetes assessment algorithms [8–11, 25–31] have been developed and some of them are especially designed for Asian [26, 31] or Chinese [8, 10, 25, 28–30], the scoring systems in most of the risk recognition models are based on the magnitude of
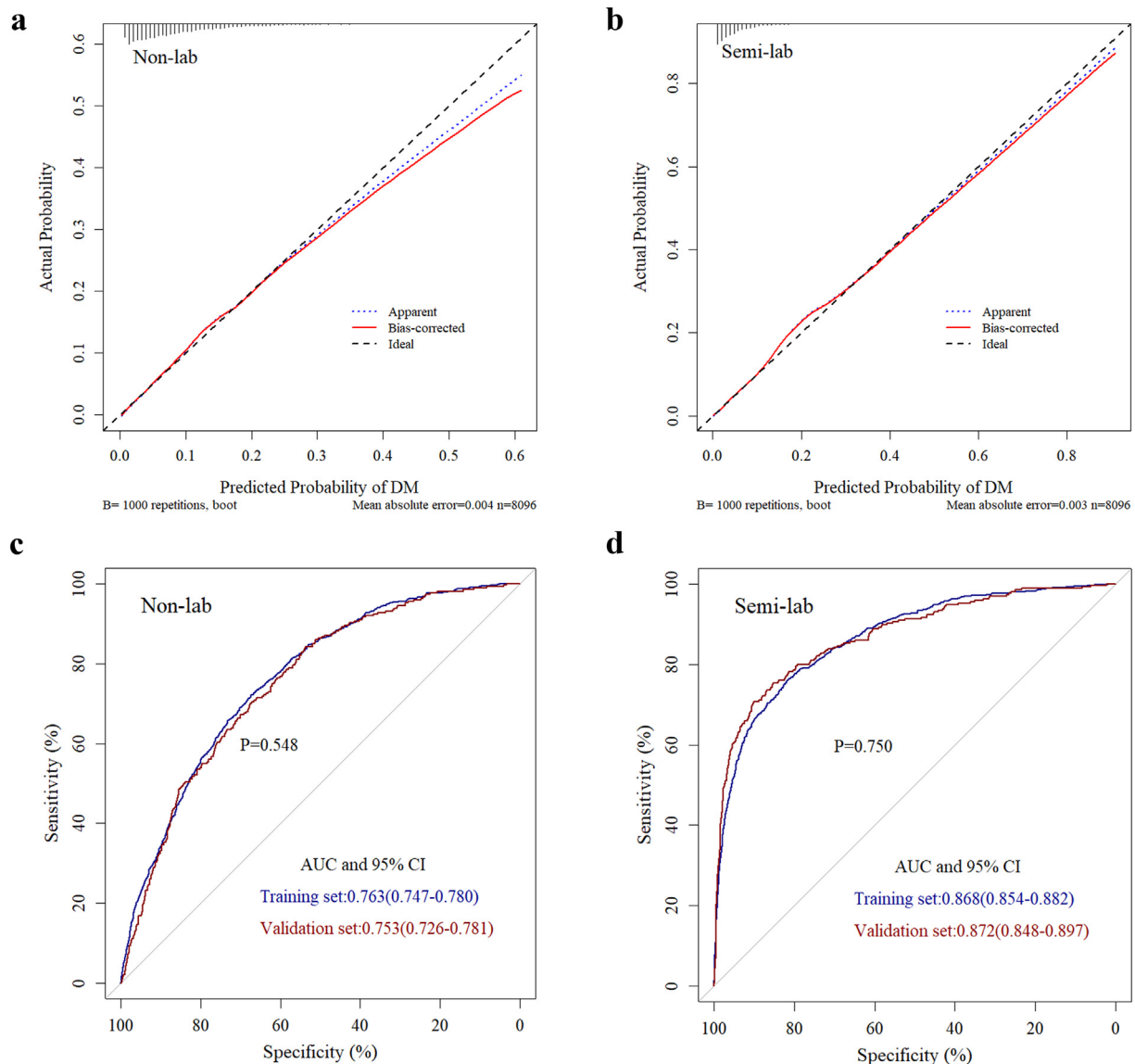
Non-lab
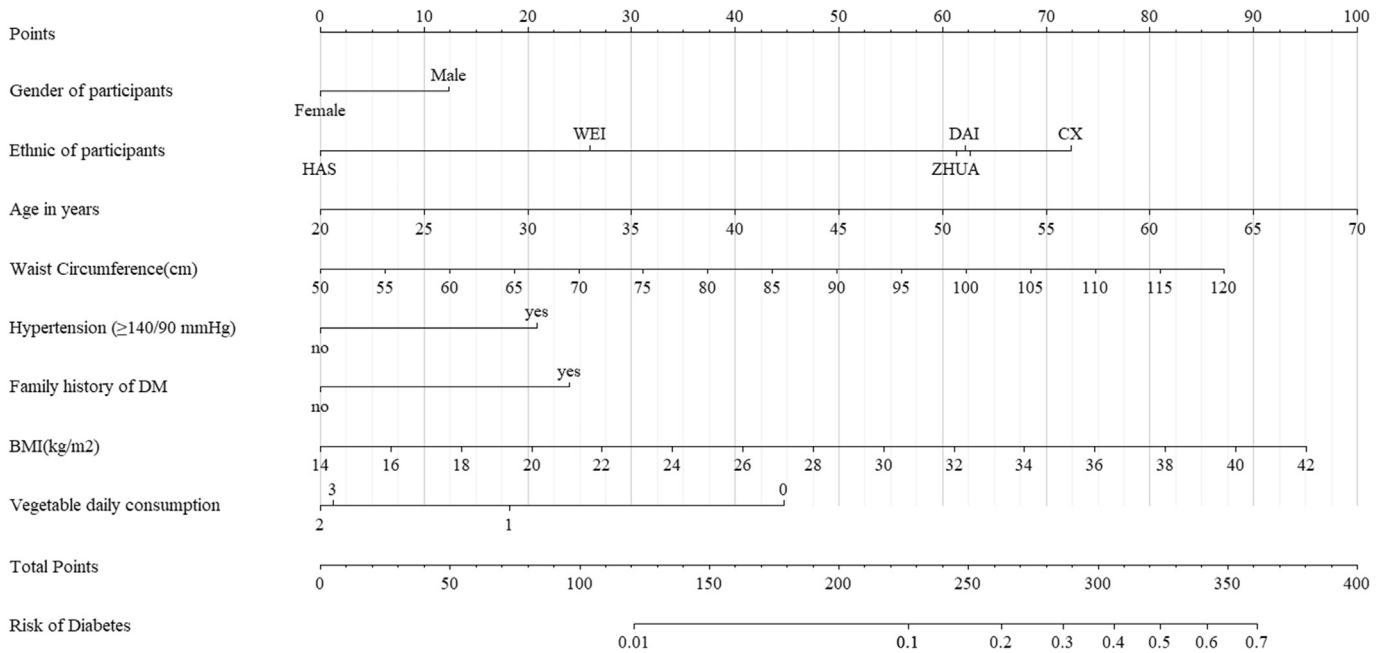
Semi-lab
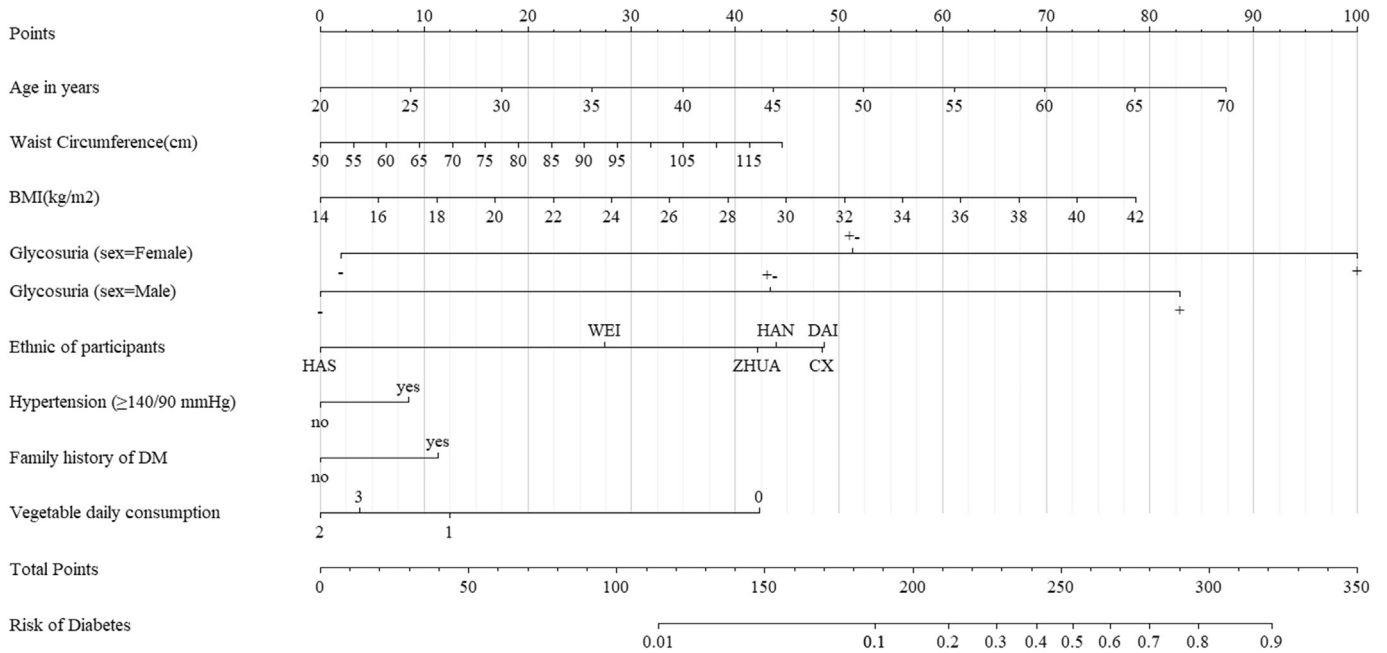
Fig. 2. Nomogram for the non-lab model and semi-lab model. HAS = The Kazak nationality. WEI = The Uyghur nationality. ZHUA = The Zhuang nationality. CX = The Korean nationality. HAN = The Han nationality. DAI = The Dai nationality. Vegetable daily consumption is a self-report variable provided from the investigated subjects. 0, 1, 2, 3 separately means very low, low, normal and high daily consumption of vegetables.

their regression coefficients [8, 11, 25, 29–31]. In this system, continuous variables are often converted into categorical variables, thus, two separate individuals with slight changes in the continuous variables (still in the same category) will get the same risk score. Besides, this scoring system generally relies on specific tables which might be inconvenient when applied in self-assessment for high risk individuals. In all the published models, only 3 used nomograms to assess the probability

of undiagnosed diabetes. Yet one [25] still adopted the scoring system as mentioned above, while the other two [9, 26] employed the more advanced Bayesian model average method to construct the nomograms. However, these two models are not fully calibrated or validated. Namely, neither these two research compared the predicted probability with the actual observed probability nor did they cut their samples into training set and validation set and validate the result in the validation

Fig. 3. Validation of non-lab model and semi-lab model. Internal validation of non-lab nomogram(a) and semi-lab nomogram(b) using the bootstrap sampling method; External validation using the receiver operating characteristic curve both in training set and validation set for non-lab nomogram(c) and semi-lab nomogram(d).

set. In addition, none of the existing models was assessed by the decision curve analysis for their clinical usefulness. In our study, the nomograms were constructed using the method mentioned above, and two websites were created with the algorithms, which would be very convenient for clinicians or individuals to assess the probability of diabetes with just a few clicks of the mouse. Both risk models were validated internally using the bootstrap sampling method and externally in the validation set. Besides, clinical usefulness was presented by the reduction in avoidable OGTT per 100 patients without increasing the number of false-positive results using the decision curve analysis method.

The non-lab model showed no significant improvement compared with the previously published model [8] because of the control of similar risk factors. However, after the introduction of the postprandial 2-h period glycosuria qualitative to the non-lab model, a significantly increment of C-index occurred. Urine glucose test for screening of diabetes is not recommended by the World Health Organization (WHO) [32] and some researchers [33] because of its low sensitivity, but they highlighted the high specificity of urine glucose test. Besides, participants in these studies were not asked to empty their bladder prior

to the examination of postprandial urine and the test was performed either 1–2 h after a solid morning or evening meal. Meanwhile, increasing evidence suggests comparable efficiencies of self-monitoring of urine glucose with self-monitoring of blood glucose on glycemic control in type 2 diabetes [34, 35]. However, repeatable application should be developed. Therefore, postprandial glycosuria qualitative, which can reflect the average level of a short time postprandial blood glucose, was proposed and added as a factor to the non-lab model by our group. On the one hand, it is quite cheap, non-invasive and easy for the self-assessment of the instructed individuals. On the other hand, the high specificity of urine glucose test can help to compromise the relatively low specificity of the non-lab factors thus bringing enormous economic benefits without lowering the accuracy.

There are some limitations of this study. Firstly, both models were not validated in a heterogeneous population. Hence, extra validation in different studies should be performed primarily for the application of our recognition models. Secondly, the postprandial 2-h period glycosuria qualitative may be influenced by renal threshold and some medications [36]. Therefore, our semi-lab model might be not suitable for
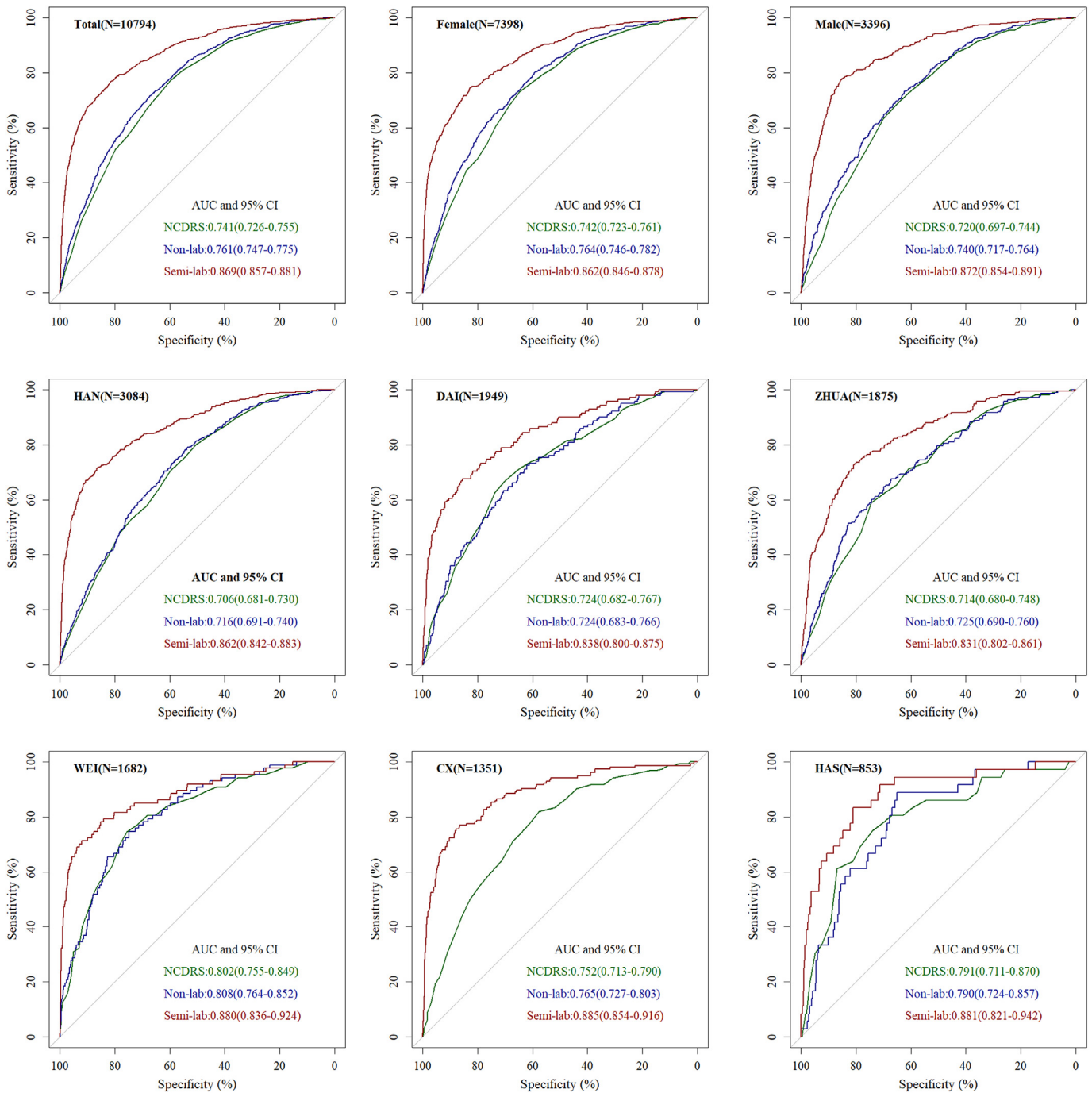
Fig. 4. Comparisons among Semi-lab, Non-lab model and The New Chinese Diabetes Risk Score in the subgroup of gender and nationality using the receiver operating characteristic curve. NCDRS = New Chinese Diabetes Risk Score. HAN = The Han nationality. CX = The Korean nationality. ZHUA = The Zhuang nationality. DAI = The Dai nationality. WEI = The Uyghur nationality. HAS = The Kazak nationality.

subjects like children, pregnant women and those who are in use of acetylsalicylic acid products or vitamin C. Thirdly, some questions and some physical examinations which were included in other famous risk scores such as "History of high blood glucose" [11], hip circumference [26] were not asked or measured. These missing variables may have a potential influence on the prediction of our models. Fourthly, there is a small proportion of subjects who refused this study. This group of subjects may have characteristics differ from the individuals who attend this study. This potential difference might influence the variable selection process to some extent. Besides, there were more women than men in this study and the sample size of the Kazak population was

much smaller than other ethnic groups. These factors may cause selection bias to some extent.

In conclusion, based on the full understanding of these limitations, our non-lab and semi-lab model showed an adequate performance for screening diabetes in different ethnic groups of China in a cost-effective way.

Considering the strength and the limitations of this present study, there is still room for improvements in the future studies. For example, new indicators of anthropometric measurement such as body volume indicator (BVI) which can reflect the volume of visceral fat measured by an iPad app (base on a 3D image technology) may be a new optional
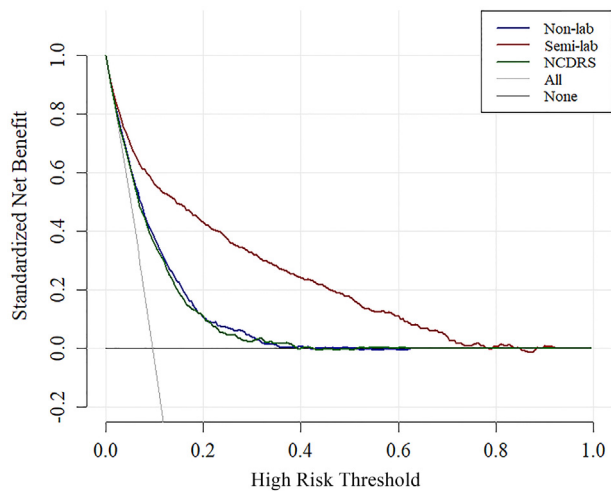
Fig. 5. Decision curve analysis for the Semi-lab, Non-lab and New Chinese Diabetes Risk Score models.

variable for the construction of the recognition model. Besides, new algorithms including neural network are also worth being employed for the detection of undiagnosed diabetes.

Author contributions

Zilin Sun, Zhiyao Wang, Wei Li and Bo Xie were responsible for the study design. Zilin Sun, Xinling Wang, Jimilanmu Maimaitiming, Xin Nian, Hong Li, Qingyun Chen, Min Liang, Qing Wang, Yan Wen, Ping Tu, Jiang Liu, Lihui Zhang, Mian Wang, Kaili Li and Li Ma were responsible for the recruitment of the participants, data acquisition. Wei Li, and Bo Xie were responsible for data collection, data analysis and writing the manuscript. Zilin Sun, Shanhu Qiu, Xin Huang, Juan Chen, Hang Wu, Xuyi Wang, Xiaohang Wang, Jingbao Liu and Min Cai were responsible for data collection and assisted with data interpretation and modifying the manuscript. Zilin Sun, Lin Guo and Fangqun Chen were responsible for the quality control of the study. Monica Sandberg and Per-Ola Carlsson were responsible for manuscript modification and discussion of the data analysis. All the authors have critically read the manuscript and approved the submitted version.

Conflicts of interest

We declare that we have no conflicts of interest.

References

[1] International Diabetes Federation, 2015. IDF diabetes atlas. 7 ed. International Diabetes Federation, Brussels, Belgium.
[2] Yang, W., Lu, J., Weng, J., et al., 2010. Prevalence of diabetes among men and women in China. N Engl J Med 362 (12), 1090–1101.
[3] Xu, Y., Wang, L., He, J., et al., 2013. Prevalence and control of diabetes in Chinese adults. JAMA 310 (9), 948–959.
[4] Wang, L., Gao, P., Zhang, M., et al., 2017. Prevalence and ethnic pattern of diabetes and prediabetes in China in 2013. JAMA 317 (24), 2515–2523.
[5] UK Prospective Diabetes Study 6, 1990. Complications in newly diagnosed type 2 diabetic patients and their association with different clinical and biochemical risk factors. Diabet Res Edinb Scotl 13 (1), 1–11.
[6] American Diabetes Association, 2017. Standards of medical care in diabetes–2017. Diabetes Care 40 (Suppl.1), S11–S25.
[7] Qiao, Q., Nakagami, T., Tuomilehto, J., et al., 2000. Comparison of the fasting and the 2-h glucose criteria for diabetes in different Asian cohorts. Diabetologia 43 (12), 1470–1475.
[8] Zhou, X., Qiao, Q., Ji, L., et al., 2013. Nonlaboratory-based risk assessment algorithm for undiagnosed type 2 diabetes developed on a nation-wide diabetes survey. Diabetes Care 36 (12), 3944–3952.
[9] Pongchaiyakul, C., Kotruchin, P., Wanothayaroj, E., Nguyen, T.V., 2011. An innovative prognostic model for predicting diabetes risk in the Thai population. Diabetes Res Clin Pract 94 (2), 193–198.
[10] Liu, M., Pan, C., Jin, M., 2011. A Chinese diabetes risk score for screening of undiagnosed diabetes and abnormal glucose tolerance. Diabetes Technol Ther 13 (5), 501–507.
[11] Lindstrom, J., Tuomilehto, J., 2003. The diabetes risk score: a practical tool to predict type 2 diabetes risk. Diabetes Care 26 (3), 725–731.
[12] Collins, G.S., Mallett, S., Omar, O., Yu, L.M., 2011. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. BMC Med 9, 103.
[13] Balachandran, V.P., Gonen, M., Smith, J.J., Dematteo, R.P., 2015. Nomograms in oncology: more than meets the eye. Lancet Oncol 16 (4), e173–e180.
[14] Vickers, A.J., Elkin, E.B., 2006. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 26 (6), 565–574.
[15] Rousson, V., Zumbrunn, T., 2011. Decision curve analysis revisited: overall net benefit, relationships to ROC curve analysis, and application to case-control studies. BMC Med Inform Decis Mak 11, 45.
[16] Zhang, Z., Kattan, M.W., 2017. Drawing nomograms with R: applications to categorical outcome and survival data. Ann Transl Med 5 (10), 211.
[17] Tang, L.Q., Li, C.F., Li, J., et al., 2016. Establishment and validation of prognostic nomograms for endemic nasopharyngeal carcinoma. J Natl Cancer Inst 108 (1).
[18] Jin, X., Jiang, Y.Z., Chen, S., Shao, Z.M., Di, G.H., 2016. A nomogram for predicting the pathological response of axillary lymph node metastasis in breast Cancer patients. Sci Rep 6, 32585.
[19] Dongsheng Yang. Build prognostic nomograms for risk assessment using SAS.In Proc SAS Global Forum 2013. Paper 264–2013. http://support.sas.com/resources/papers/proceedings13/264–2013.pdf, 196KB. Accessed February 22, 2014.
[20] Delong, E.R., Delong, D.M., Clarke-Pearson, D.L., 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44 (3), 837–845.
[21] Hijazi, Z., Oldgren, J., Lindbäck, J., et al., 2016. The novel biomarker-based ABC (age, biomarkers, clinical history)-bleeding risk score for patients with atrial fibrillation: a derivation and validation study. Lancet 387 (10035), 2302–2311.
[22] Chiu, M., Austin, P.C., Manuel, D.G., Shah, B.R., Tu, J.V., 2011. Deriving ethnic-specific BMI cutoff points for assessing diabetes risk. Diabetes Care 34 (8), 1741–1748.
[23] Feng, Y., Yang, Y., Ma, X., et al., 2012. Prevalence of diabetes among Han, Manchu and Korean ethnicities in the Mudanjiang area of China: a cross-sectional survey. BMC Public Health 12, 23.
[24] Tao, Y., Mao, X., Xie, Z., 2008. The prevalence of type 2 diabetes and hypertension in Uygur and Kazak populations. Cardiovasc Toxicol 8 (4), 155–159.
[25] Wong, C.K., Siu, S.C., Wan, E.Y., et al., 2016. Simple non-laboratory- and laboratory-based risk assessment algorithms and nomogram for detecting undiagnosed diabetes mellitus. J Diabetes 8 (3), 414–421.
[26] Ta, M.T., Nguyen, K.T., Nguyen, N.D., Campbell, L.V., Nguyen, T.V., 2010. Identification of undiagnosed type 2 diabetes by systolic blood pressure and waist-to-hip ratio. Diabetologia 53 (10), 2139–2146.
[27] Kengne, A.P., Beulens, J.W., Peelen, L.M., et al., 2014. Non-invasive risk scores for prediction of type 2 diabetes (EPIC-InterAct): a validation of existing models. Lancet Diabet Endocrinol 2 (1), 19–29.
[28] Xie, J., Hu, D., Yu, D., Chen, C.-S., He, J., Gu, D., 2010. A quick self-assessment tool to identify individuals at high risk of type 2 diabetes in the Chinese general population. J Epidemiol Community Health 64 (3), 236–242.
[29] Ko, G., So, W., Tong, P., et al., 2010. A simple risk score to identify southern Chinese at high risk for diabetes. Diabet Med 27 (6), 644–649.
[30] Gao, W., Dong, Y., Pang, Z., et al., 2010. A simple Chinese risk score for undiagnosed diabetes. Diabet Med 27 (3), 274–281.
[31] Aekplakorn, W., Bunnag, P., Woodward, M., et al., 2006. A risk score for predicting incident diabetes in the Thai population. Diabetes Care 29 (8), 1872–1877.
[32] World Health Organization, 2003. Screening for type 2 diabetes: report of a World Health Organization and International Diabetes Federation meeting.
[33] Friderichsen, B., Maunsbach, M., 1997. Glycosuric tests should not be employed in population screenings for NIDDM. J Public Health Med 19 (1), 55–60.
[34] Dallosso, H.M., Bodicoat, D.H., Campbell, M., et al., 2015. Self-monitoring of blood glucose versus self-monitoring of urine glucose in adults with newly diagnosed type 2 diabetes receiving structured education: a cluster randomized controlled trial. Diabet Med 32 (3), 414–422.
[35] Lu, J., Bu, R.F., Sun, Z.L., et al., 2011. Comparable efficacy of self-monitoring of quantitative urine glucose with self-monitoring of blood glucose on glycaemic control in non-insulin-treated type 2 diabetes. Diabetes Res Clin Pract 93 (2), 179–186.
[36] Goldstein, D.E., Little, R.R., Lorenz, R.A., et al., 2004. Tests of glycemia in diabetes. Diabetes Care 27 (7), 1761–1773.