의학석사 학위논문

Cross-sectional and Longitudinal Validation of Risk Prediction Models for Type 2 Diabetes Mellitus

제 2 형 당뇨병 위험 예측 모델의 단면적, 종적 예측력 검증

2015 년 2 월

서울대학교 대학원
의학과 분자유전체의학 전공
안 창 호

A thesis of the Master's degree

제 2 형 당뇨병 위험 예측 모델의 단면적, 종적 예측력 검증

Cross-sectional and Longitudinal Validation of Risk Prediction Models for Type 2 Diabetes Mellitus

February 2015

The Department of Medicine,
Seoul National University
College of Medicine
Chang Ho Ahn

제 2 형 당뇨병 위험 예측 모델의 단면적, 종적 예측력 검증

지도교수 조 영 민

이 논문을 의학석사 학위논문으로 제출함

2014 년 10 월

서울대학교 대학원

의학과 분자유전체의학 전공

안 창 호

안창호의 의학석사 학위논문을 인준함

2015 년 1 월

위 원 장 (인)

부위원장 (인)

위 원 (인)

Cross-sectional and Longitudinal Validation of Risk Prediction Models for Type 2 Diabetes Mellitus

by
Chang Ho Ahn

A thesis submitted to the Department of Medicine in partial fulfillment of the requirements for the Degree of Master of Philosophy in Molecular Genomics at Seoul National University College of Medicine

January 2015

Approved by Thesis Committee:

Professor _____Chairman
Professor _____Vice chairman
Professor _____

ABSTRACT

Introduction: Early detection of undiagnosed diabetes and prediction of future diabetes are crucial for preventing or delaying the detrimental complications of diabetes. For this purpose, various risk prediction models including a recently published Korean screening score composed of non-laboratory parameters have been developed. We evaluated the validity of the Korean screening score for undiagnosed and incident diabetes in an independent study population. Further, its predictive performance was compared with various other non-laboratory risk prediction models and laboratory parameters.

Methods: The data of 26,675 individuals who visited Seoul National University Hospital Healthcare System Gangnam Center for health screening program were reviewed for cross-sectional validation of undiagnosed diabetes and the data of 3,029 individuals with mean 6.2 years of follow-up were reviewed for longitudinal validation of incident diabetes. The predictive performance of the Korean screening score, other 16 previously published risk prediction models and the risk prediction model of laboratory parameters were compared.

Results: For the screening of undiagnosed diabetes, the Korean screening score exhibited a sensitivity of 81%, a specificity of 58%, and an area under the curve of receiver operating characteristic curve (AROC) of 0.754. All the other non-laboratory risk prediction models revealed comparable AROC. For the prediction of incident diabetes, the Korean score demonstrated a

sensitivity of 74%, a specificity of 54%, and an AROC of 0.696. Relative to the Korean score, the risk prediction by the laboratory parameters - fasting plasma glucose (FPG) and hemoglobin A1c (HbA1c) level - demonstrated a significantly higher AROC (0.838 vs. 0.696, P value <0.001). Combining of FPG, HbA1c and the Korean score data increased AROC by small increment (0.838 vs. 0.849, P value = 0.016) without statistically significant improvement in risk classification (net reclassification index 4.6%, P value = 0.264, integrated discrimination improvement 0.006, P value = 0.006).

Conclusions: In conclusion, the Korean screening score is useful for detecting undiagnosed diabetes but inferior to the laboratory parameters for the prediction of incident diabetes.

Keywords: diabetes mellitus, risk assessment, validation studies

Student number: 2013-21733

CONTENTS

LIST OF TABLES AND FIGURES

INTRODUCTION

The burden of diabetes mellitus is constantly increasing worldwide. International Federation of Diabetes estimated that globally 382 million people have diabetes in 2013 and this number will be increased by 55% until 2035 (1). Since late 1990s, large randomized controlled trials demonstrated that intensive glycemic control can reduce the progression of microvascular complications of diabetes for diabetic patients (2, 3), and lifestyle and pharmacological interventions can delay or prevent type 2 diabetes for prediabetic patients (4, 5). These results were further confirmed by meta-analysis (6-8). Therefore, identification of individuals with undiagnosed diabetes or high risk of future diabetes is of paramount importance to fight the global epidemic of diabetes. However significant proportions of diabetes patients are unaware of their conditions and left untreated. It was estimated that 27.8% of diabetes patients in the United States (9) and 27.3% in Korea were left undiagnosed (10). To identify this population with potentially modifiable health outcome, an effective screening program is essential.

For this purpose, various risk prediction models were constructed (11-13). Each risk prediction model was designed to identify either undiagnosed diabetes, incident diabetes, or both (11-13). The development of the risk prediction models was based on various populations with distinct ethnic and medical background (11-13). Known risk factors for diabetes, such as central obesity, family history of diabetes, and old age were included in the majority of the models (12). Other potential risk factors, such as steroid use, intake of

red meat, coffee consumption, or alcohol consumption were also included in some of the models (12). The differences in the primary end-point of the model, the population that the model was derived from, and risk predictors included in the model lead to considerable heterogeneity among the risk prediction models and great caution should be made before their application to a new population.

The risk prediction models can be categorized into the models solely based on non-laboratory parameters and the others based on both laboratory and non-laboratory parameters (11). These two types of risk prediction models need different amount of resources and have different range of applicability. The models with non-laboratory parameters have virtually no cost and are ready to be used by a lay person. Therefore, they are suitable for whole population based screening and can be applied to both screening of undiagnosed diabetes and prediction of incident diabetes. The models using laboratory parameters need additional cost and time, which limit their use in the population based screening (14). In addition, screening undiagnosed diabetes with laboratory parameters has practically little impact because a simple fasting blood test can confirm the diagnosis of diabetes. Therefore, the models using laboratory parameters are better applied to the prediction of incident diabetes. Generally, the use of laboratory parameters improves the discriminatory ability of risk prediction models, but by different degrees in each study (14, 15). In Korea, the National Health Insurance Program provides biannual health screening program including fasting blood test for all adults with age over 40 (16). Therefore, the results of laboratory tests, especially basic blood tests, are

readily available in Korea. However the risk prediction of diabetes using laboratory parameters and non-laboratory parameters has not been thoroughly evaluated and compared yet.

Recently, a simple Korean screening score for diabetes composed of non-laboratory parameters have been developed based on the result of the Korean National Health and Nutrition Examination Survey (KNHANES) (17). The Korean score developed from the data of 2001 and 2005 KNHANES survey and validated in the data of 2007 and 2008 KNHANES survey. It was originally designed to screen undiagnosed diabetes and scored the risk of undiagnosed diabetes by age, family history of diabetes, hypertension, waist circumference, smoking status and alcohol intake from 0 to 11. The estimated risk of undiagnosed diabetes was 2% for score 0 to 4, 6% for score 5 to 7, 12% for score 8 to 9, 19% for score 10 to 11. The cut-off for the high risk of diabetes was determined as 5 or more which resulted in the highest Youden index (sensitivity + specificity -1)(17). Although the Korean screening score was validated in the national survey data, it has not been validated in other population data. In addition, its use for the prediction of future diabetes have not been evaluated and compared with laboratory parameters. We performed a comprehensive validation of the Korean screening score in a large independent Korean population data and compared it with various other risk prediction models composed of non-laboratory parameters for both undiagnosed and incident diabetes. Then, we compared the performance of Korean screening score and laboratory parameters for the prediction of incident diabetes.

METHODS

Subjects and study design

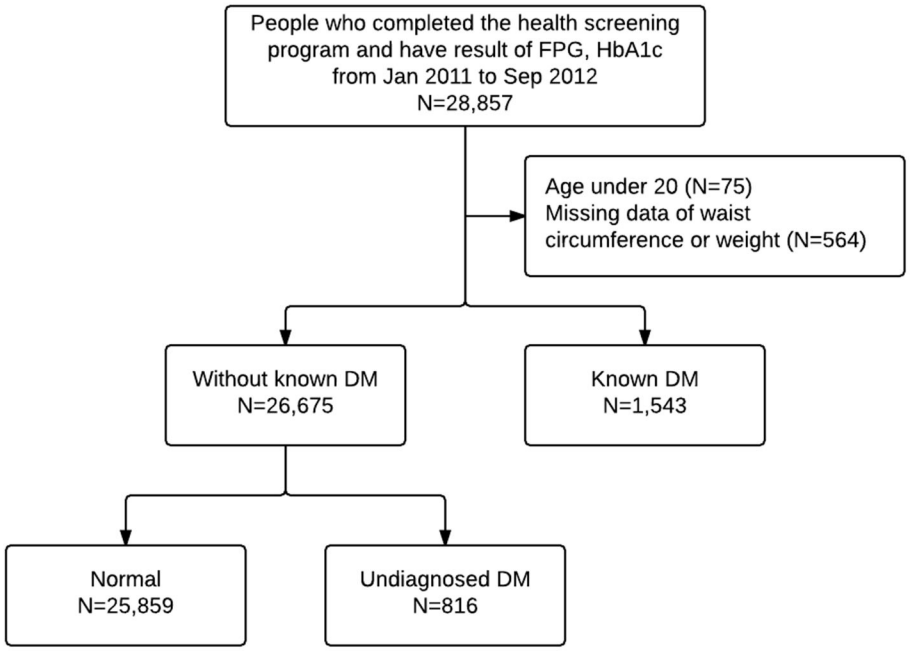
Subjects were recruited at Seoul National University Hospital Healthcare System Gangnam Center. This center provides health-screening program composed of basic and specialized examinations. The basic examinations include health-related questionnaire, anthropometric measurements, biochemical tests, abdominal ultrasonography and upper-GI endoscopy. The specialized examinations which include various imaging and functional studies were provided after the individual and physician's discussion on the necessity of the tests. The individuals were asked to provide their results of health-screening program for research and the database was constructed after encryption of the personal information. This study was approved by the Institutional Review Board of Seoul National University Hospital (IRB No. 1308-004-507).

Cross-sectional validation for screening of undiagnosed diabetes

Among the individuals who visited healthcare center from January 1st 2011 to September 31st 2012, total 28857 individuals completed the health screening program and had results of fasting plasma glucose (FPG) and hemoglobin A1c (HbA1c) level. We excluded 75 individuals under age of 20, 564 individuals with missing data of waist circumference or weight, and 1543 individuals who reported to have known diabetes. Consequently 26675 individuals comprised the study population for cross-sectional validation. In

cases of individuals who had multiple visits during the study period, only the data of first visit were used (Figure 1).

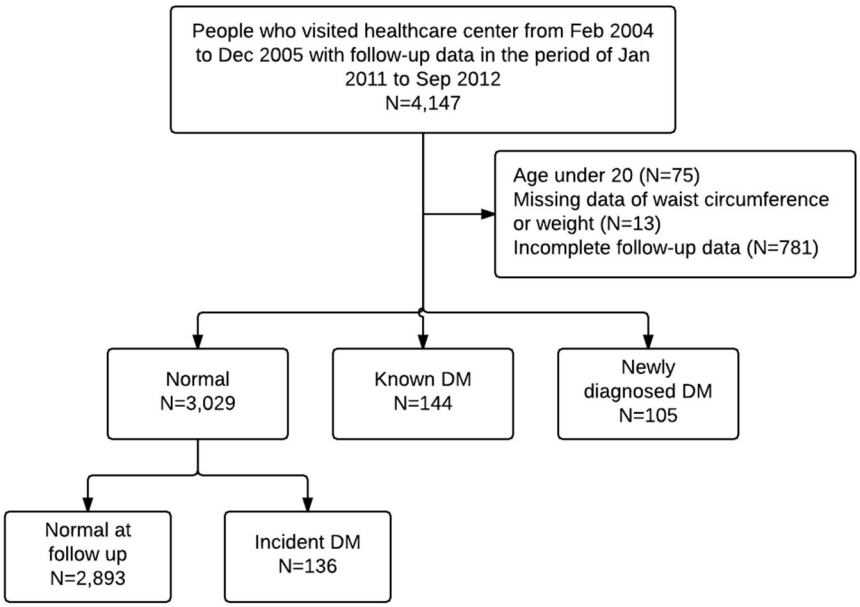
Figure 1. Study population of the cross-sectional validation



Longitudinal validation for prediction of incident diabetes

Among the individuals who visited our center from February 1st 2004 to Dec 31st 2005, 4147 individuals completed the health-screening program and had follow-up visits in the period of January 1st 2011 to September 31st 2012. From this eligible population, the 75 individuals under age of 20, 13 individuals with missing data in waist circumference or weight, and 781 individuals with incomplete follow-up data were excluded. The 144 individuals with previously diagnosed diabetes and 105 individuals with newly diagnosed diabetes at baseline visit were further excluded. As a result, 3029 individuals comprised the study population for longitudinal validation with the mean follow-up duration of 6.2 years (Figure 2).

Figure 2. Study population of longitudinal validation



Clinical and laboratory evaluation

The questionnaire gathered information on medical history, family history, health-related habits and physical activity. Medical history includes current medications and previous diagnosis of diabetes, hypertension, and dyslipidemia. Individuals who reported to have hypertension, were taking antihypertensive medication or whose systolic blood pressure was $\geq$140/90 mmHg were defined to have hypertension. Family history of diabetes was confined to first degree relatives. Smoking status was classified as current smoker (who is currently smoking), ex-smoker (who is not currently smoking but had smoked at least 5 packs of cigarettes in the lifetime), and never smoker. Alcohol consumption was calculated as average daily number of drinks based on the frequency of drinking per week and the amount of alcohol beverage consumed. Physical activity was assessed based on leisure time physical activity. Subjects were classified as physically inactive if they do any levels of activity no more than 10 minutes per week. The blood samples were collected after 12 hours' overnight fast. Plasma glucose, HbA1c, total cholesterol, HDL cholesterol and triglyceride level were measured.

Definition of diabetes mellitus

The subjects who answered "Yes" to the question, "Have you ever been diagnosed with diabetes by a physician?" or were taking anti-diabetic medications were defined as having 'known diabetes'. The subjects who were first diagnosed with type 2 diabetes based on the result of the fasting blood test at the health-screening program were classified as 'undiagnosed diabetes'.

The incident diabetes was defined as the subjects who initially did not have known or undiagnosed diabetes at the baseline visit of the longitudinal study and became known or undiagnosed diabetes at the follow-up visit. Diabetes was diagnosed in subjects with FPG ≥126 mg/dl or HbA1c ≥6.5% according to the 2010 revision of the American Diabetes Association (ADA) guidelines (18).

Risk prediction models

We adopted risk prediction models from three recently published systematic reviews on risk prediction models for type 2 diabetes (11-13). According to the standard methodology for systematic review, each systematic review searched published articles which reported risk prediction models for type 2 diabetes. The timings of the searches in the three reviews were January 2011, February 2011, and May 2011, respectively. We further searched other studies reporting risk prediction models published from January 2011 to May 2013 on PubMed and Google Scholar using the following search string: ((“diabetes” OR “type 2 diabetes”) AND (“score” OR “model” OR “prediction”)). Among the identified risk prediction models, the models using only non-laboratory parameters were included. When the definition of each variable was not identical with our study, we tried to use the best available variable. Two variables, the current use of corticosteroid and history of gestational diabetes, were omitted because those conditions were not collected in our study.

Statistical analysis

We classified subjects according to their diabetes status for descriptive statistics. Continuous variables were expressed as means ± SD and categorical variables were presented as percentages. The difference between each group was analyzed by t test or chi-square test. To validate and compare various risk prediction models, we calculated the proportion of high risk individuals, sensitivity, specificity, positive predictive value, negative predictive value, likelihood ratios and Youden index (sensitivity + specificity -1) for each model. Area under the receiver operating characteristic curve (AROC) was also calculated as a discrimination index. In the study population for cross-sectional validation, we applied risk prediction models to detect undiagnosed diabetes. In the study population for longitudinal validation, we applied risk scores to the non-diabetic subjects at the baseline visit to predict incident diabetes at follow-up visit.

To compare the performance of non-laboratory risk prediction model and laboratory parameters for incident diabetes, we applied multivariate logistic regression analysis for laboratory parameters including fasting plasma glucose, HbA1c, total cholesterol, HDL cholesterol and triglyceride level in the population of the longitudinal study. We calculated aforementioned measures including AROC, net reclassification index (NRI) and integrated discrimination improvement (IDI). Addition to this, to assess the effectiveness of each screening program, we simulated the application of the Korean screening score and combined risk prediction model of the Korean score and laboratory parameters to the baseline visit of longitudinal study (total 3134 individuals without known diabetes at baseline visit, which were 3029 normal

individuals plus 105 undiagnosed diabetes at baseline). All statistical analyses were performed using SPSS v18.0 (SPSS inc. Chicago, IL, USA) or R v3.0.1 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Characteristics of the study population for cross-sectional validation

The characteristics of the subjects who comprised the study population for cross-sectional validation according to the diabetes status were summarized in table 1. The prevalence of the diabetes was 8.4% including 2.9% of undiagnosed diabetes. Known risk factors including old age, higher BMI, waist circumference, hypertension history, family history of diabetes and smoking were associated with diabetes. The two groups, known and undiagnosed diabetes had similar fasting blood glucose level (132.4 ±35.3 vs. 131.3 ±25.9, P value = 0.396), but the total cholesterol and triglyceride level were lower in known diabetes. The modifiable risk factors of diabetes (obesity, physical activity and alcohol intake) were significantly different between known and undiagnosed diabetes. Known diabetes was less obese (BMI 24.6 ±3.0 vs. 25.6 ±3.2, P value <0.001), does more regular exercise (81.3% vs. 69.7%, P value <0.001), and had less alcohol intake (1.6 ±1.9 vs. 1.9 ±2.1, P value = 0.014) than undiagnosed diabetes.

Table 1. Baseline characteristics of the study population for cross-sectional validation

Characteristics	Normal n = 25,859	Undiagnosed DM n = 816	Known DM n = 1,543	P value Normal vs. undiagnosed DM	P value Normal vs. known DM	P value Undiagnosed DM vs. known DM
Age (years)	48.5 ±11.1	55.6 ±9.5	58.1 ±9.8	<0.001	<0.001	<0.001
Male	13,890 (53.7)	592 (72.5)	1,172 (76)	<0.001	<0.001	0.070
Height (cm)	165.7 ±8.2	166.9 ±8.3	166.4 ±7.8	<0.001	0.001	0.235
Weight (kg)	63.9 ±11.9	71.5 ±11.8	68.4 ±11.1	<0.001	<0.001	<0.001
BMI (kg/m^2)	23.1 ±3.1	25.6 ±3.2	24.6 ±3.0	<0.001	<0.001	<0.001
Waist circumference (cm)	83.4 ±10.3	90.6 ±8.1	88.6 ±8.2	<0.001	<0.001	<0.001
Hypertension history	5,705 (22.1)	407 (49.9)	841 (54.5)	<0.001	<0.001	0.032
Family history of diabetes	5,015 (19.4)	258 (31.6)	649 (42.1)	<0.001	<0.001	<0.001
Current Smoking	4,330 (16.7)	172 (21.1)	323 (20.9)	0.001	<0.001	0.934
Regular exercise	19,991 (77.3)	569 (69.7)	1,254 (81.3)	<0.001	<0.001	<0.001
Alcohol intake (glasses/day)	1.3 ±1.8	1.9 ±2.1	1.6 ±1.9	<0.001	<0.001	0.014
Fasting blood glucose (mg/dL)	93.5 ±9.8	131.3 ±25.9	132.4 ±35.3	<0.001	<0.001	0.396
HbA1c (%)	5.6 ±0.3	6.8 ±0.9	6.9 ±1.1	<0.001	<0.001	0.006
Total cholesterol (mg/dL)	194.1 ±33.1	201 ±40.4	175.9 ±36.2	<0.001	<0.001	<0.001
Triglyceride (mg/dL)	107.6 ±73.8	157 ±96.9	130.4 ±90.6	<0.001	<0.001	<0.001
HDL cholesterol (mg/dL)	53 ±11.6	48.2 ±9.9	48.3 ±10.5	<0.001	<0.001	0.860

Abbreviations: DM, diabetes mellitus; BMI, body mass index; HbA1c, hemoglobin A1c.

Data are mean ±SD or n (%).

Characteristics of the study population for longitudinal validation

The baseline characteristics of the subjects who comprised the study population for longitudinal validation according to the diabetes status were summarized in table 2. The 4.5% of the subjects who were normoglycemic at baseline developed incident diabetes between baseline and follow-up visits. The minimum, maximum and mean duration of follow-up were 5.0, 7.9 and 6.2 years. Among various parameters, old age, high BMI, waist circumference, hypertension history and family history of diabetes were associated with incident diabetes. The proportion of current smoker and the amount of daily alcohol intake were not significantly different between normal and incident diabetes at follow-up.

Table 2. Baseline characteristics of the study population for longitudinal validation

Characteristics	Normal at f/u n = 2,893	Incident DM n = 136	P value
Age (years)	47.1 ±9.8	52.0 ±8.2	<0.001
Male	1659 (57.3)	95 (69.9)	0.004
Height (cm)	165.7 ±8.0	166.8 ±8.0	0.108
Weight (kg)	64.2 ±11.1	70.6 ±11.3	<0.001
BMI (kg/m^2)	23.3 ±2.8	25.3 ±2.9	<0.001
Waist circumference (cm)	84.0 ±7.7	89.3 ±7.0	<0.001
Hypertension history	743 (25.7)	56 (41.2)	<0.001
Family history of diabetes	610 (21.1)	44 (32.4)	0.002
Current smoking	528 (18.3)	32 (23.5)	0.121
Regular exercise	1767 (61.1)	88 (64.7)	0.396
Alcohol intake (glasses/day)	1.1 ±1.9	1.4 ±2.0	0.079
Fasting blood glucose (mg/dL)	96.1 ±9.8	106.8 ±10.7	<0.001
HbA1c (%)	5.5 ±0.3	5.9 ±0.3	<0.001
Total cholesterol (mg/dL)	197.1 ±32.8	206.0 ±32.6	0.002
Triglyceride (mg/dL)	108.6 ±69.4	144.5 ±80.7	<0.001
HDL cholesterol (mg/dL)	54.7 ±13.2	50.1 ±13.3	<0.001

Abbreviations: DM, diabetes mellitus; BMI, body mass index; HbA1c, hemoglobin A1c.

Data are mean ±SD or n (%).

Cross-sectional validation for screening of undiagnosed diabetes

We evaluated the performance of different risk prediction models for the screening of undiagnosed diabetes with each model's original cut-off value and new cut-off value showing the highest Youden index (Table 3, Figure 1). The Korean score (17) demonstrated an AROC of 0.754 (0.740-0.769), a sensitivity of 91% and a specificity of 40% with original cut-off value. The sensitivity and specificity of the Korean score were 81% and 58% with new cut-off value

The other 16 risk prediction models exhibited AROCs ranged from 0.697 to 0.782. Among them, Australian score (19), Danish score (20), ADA questionnaire (21), Japanese score (TOPICS-10 study) (22) and the Leicester Risk Assessment score (23) had significantly higher AROC than the Korean score. The 15 models including the Korean score needed readjustment of cut-off values in our cohort. With these new cut-off values, the sensitivities of the scores were varied from 68 to 85% and the specificities from 42 to 72%. The models classified 37.3% (minimum 29%, maximum 59%) of subjects as high risk of having undiagnosed diabetes.

Table 3. Performance of risk prediction models in cross-sectional validation for screening of undiagnosed diabetes

Risk prediction model	Cut-off	Patients at high risk (%)	AROC	95% CI	Sensitivity (%)	Specificity (%)	Youden Index	PPV (%)	NPV (%)	LR+	LR-	P value*
Korean score (17)	≥6	43	0.754	0.740-0.769	81	58	40	5.8	99	1.95	0.32	
	≥5	61			91	40	32	4.6	99	1.53	0.22	
Australian score (AUSDRISK study) (19)	≥13	29	0.782	0.769-0.796	70	72	43	7.4	99	2.54	0.41	<0.001
	≥12	35			75	66	41	6.6	99	2.22	0.38	
Danish score (20)	≥25	36	0.777	0.763-0.792	77	65	43	6.6	99	2.23	0.35	0.001
	≥31	20			55	81	36	8.4	98	2.90	0.55	
ADA questionnaire (21)	≥3	36	0.776	0.762-0.790	78	65	44	6.6	99	2.25	0.33	0.004
	≥5	6			21	95	16	11.5	97	4.14	0.83	
Japanese score (TOPICS-10 study) (22)	≥8**	32	0.774	0.760-0.788	71	70	41	6.9	99	2.33	0.42	0.002
The Leiscester Risk Assessment score (23)	≥17	30	0.773	0.759-0.788	71	71	42	7.2	99	2.46	0.41	0.007
	≥16	38			78	64	42	6.3	99	2.14	0.35	
Thai score (34)	≥8	34	0.763	0.749-0.778	72	68	39	6.5	99	2.22	0.42	0.114
	≥7	46			83	55	38	5.5	99	1.84	0.32	
Finnish score (DETECT-2 study) (26)	≥5	38	0.759	0.744-0.774	75	63	38	6.0	99	2.03	0.40	0.491
	≥7	19			50	82	32	8.2	98	2.82	0.61	
Brazilian score (35)	≥11	46	0.751	0.737-0.766	85	55	40	5.7	99	1.90	0.27	0.723
	≥18	16			42	85	27	8.1	98	2.78	0.69	
Indian score (36)	≥18	35	0.751	0.736-0.766	74	66	40	6.5	99	2.19	0.39	0.617
	≥17	39			77	63	39	6.1	99	2.05	0.37	

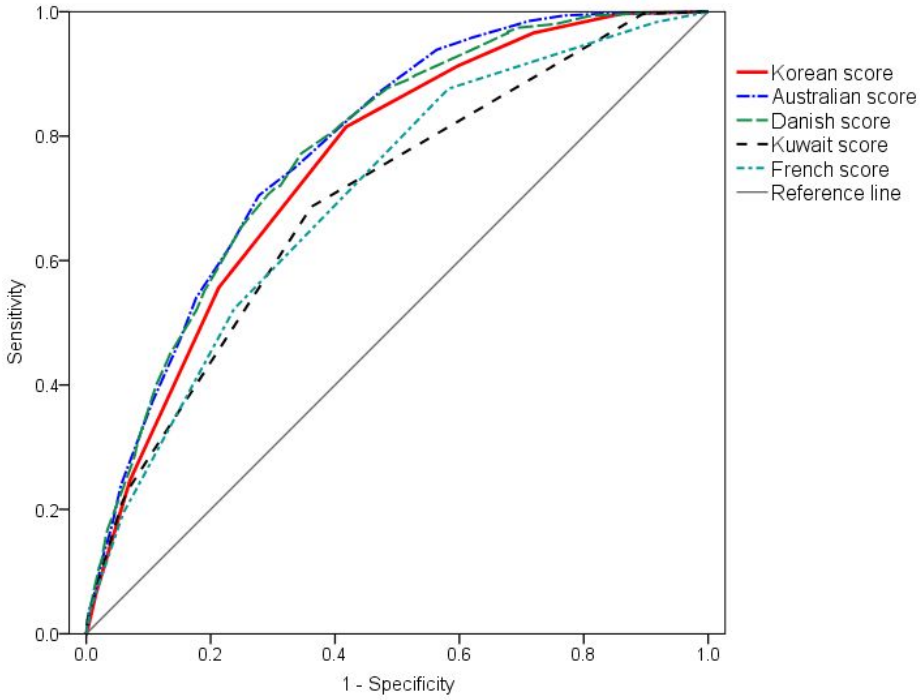
Japanese score (Ninomiya et al.) (15)	≥12	32	0.750	0.734-0.767	70	69	39	6.6	99	2.25	0.44	0.552
	≥14	23			56	79	35	7.7	98	2.63	0.55	
Chinese score (37)	≥16	40	0.733	0.717-0.748	72	61	33	5.5	99	1.85	0.46	0.004
	≥14	62			92	39	32	4.6	99	1.52	0.20	
British score (14)	≥4	39	0.730	0.714-0.747	74	62	36	5.8	99	1.95	0.42	0.002
	≥6	15			41	86	27	8.5	98	2.94	0.68	
Rotterdam model (38)	≥33.9	35	0.727	0.710-0.743	68	67	35	6.0	99	2.04	0.48	0.001
	≥37	23			52	78	30	6.9	98	2.35	0.62	
Oman score (39)	≥11	33	0.726	0.710-0.742	68	68	35	6.2	99	2.10	0.48	<0.001
	≥10	36			70	65	35	5.9	99	2.00	0.46	
French score (DESIR study) (29)	≥2**	59	0.705	0.688-0.721	88	42	29	4.5	99	1.50	0.30	<0.001
Kuwait score (40)	≥19	37	0.697	0.681-0.714	69	64	32	5.6	98	1.89	0.49	<0.001
	≥32	7			23	93	16	9.8	97	3.43	0.82	

Abbreviations: AROC, area under the curve of receiver operating characteristic curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; LR+, positive likelihood ratio; LR-, negative likelihood ratio.

Risk prediction models other than Korean score were arranged in the order of higher AROC value. The performance of different risk prediction models for screening of undiagnosed diabetes were evaluated with each model's original cut-off value and new cut-off value showing the highest Youden index. The results in the upper row of each score are based on the new cut-off value with the highest Youden

index. The results in the bottom row of each score are based on the original cut-off value. *P value for the comparison of ROC curves between Korean score and other models. **The new cut-off and original cut-off were same for Japanese score and French score.

Figure 3. ROC curves of risk prediction models in cross-sectional validation.



ROC curves of the Korean score, the 2 scores with the highest AROC, and the 2 scores with the lowest AROC.

Longitudinal validation for prediction of incident diabetes

The comparison of the performance of different risk prediction models for the prediction of incident diabetes was summarized in the table 4 (Figure 2). All risk prediction models demonstrated lower values of AROC in longitudinal validation than in cross-sectional validation. The Korean score demonstrated an AROC of 0.696 (0.656-0.736). The sensitivity and specificity of the Korean score were 89% and 37% with original cut-off value and 74% and 54% with new cut-off value determined by the highest Youden index. The other 16 models exhibited AROCs ranged from 0.630 to 0.721. The 15 out of 17 models needed readjustment of cut off value. With the new cut-off values, the models had sensitivities ranged from 51 to 86% and specificities ranged from 45 to 68%. In average, 42.5% (minimum 26%, maximum 57%) of subjects were classified as high risk of developing incident diabetes by the risk prediction models.

Table 4. Performance of risk prediction models in longitudinal validation for prediction of incident diabetes

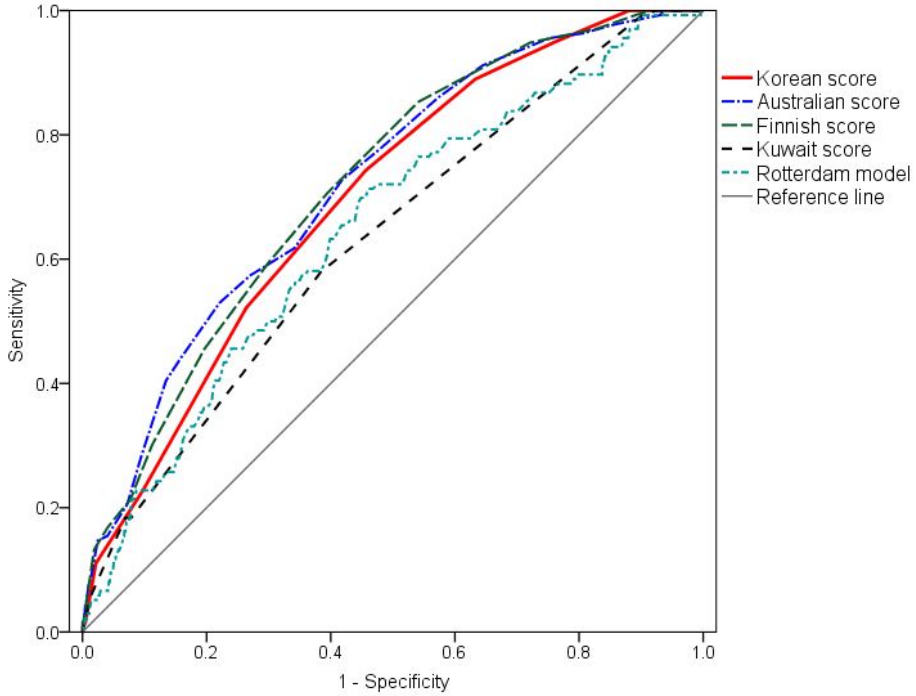
Risk prediction models	Cut-off	Patients at high risk(%)	AROC	95% CI	Sensitivity (%)	Specificity (%)	Youden Index	PPV (%)	NPV (%)	LR+	LR-	P value*
Korean score (17)	≥6	47	0.696	0.656-0.736	74	54	29	7.1	98	1.63	0.47	
	≥5	64			89	37	26	6.2	99	1.41	0.30	
Australian score (AUSDRISK study) (19)	≥11	43	0.721	0.679-0.762	73	58	31	7.6	98	1.74	0.47	0.121
	≥12	35			62	66	28	7.8	97	1.81	0.58	
Finnish score (DETECT-2 study) (26)	≥4	55	0.718	0.678-0.758	85	46	31	6.9	99	1.58	0.32	0.205
	≥7	21			46	80	26	9.8	97	2.32	0.68	
Thai score (34)	≥6	56	0.713	0.675-0.752	86	45	31	6.9	99	1.58	0.31	0.309
	≥7	47			76	54	30	7.2	98	1.66	0.44	
Danish score (20)	≥25	38	0.700	0.658-0.742	65	64	29	7.8	98	1.80	0.54	0.845
	≥31	22			43	79	22	8.8	97	2.06	0.72	
The Leiscester Risk Assessment score (23)	≥13	45	0.697	0.655-0.739	74	57	30	7.4	98	1.70	0.47	0.967
	≥16	35			62	66	28	8.0	97	1.84	0.58	
Japanese score (TOPICS-10 study) (22)	≥8**	33	0.696	0.655-0.738	58	68	26	7.9	97	1.83	0.61	0.976
Chinese score (37)	≥16	37	0.692	0.651-0.732	66	64	31	8.0	98	1.86	0.53	0.838
	≥14	58			84	43	27	6.4	98	1.46	0.38	
Indian score (36)	≥18	38	0.689	0.648-	64	63	27	7.5	97	1.73	0.57	0.703

	≥17	42		0.729	68	59	26	7.2	97	1.64	0.55	
Japanese score (Ninomiya et al.) (15)	≥12	39	0.688	0.643-0.733	69	62	31	7.9	98	1.83	0.50	0.658
	≥14	28			57	73	30	9.0	97	2.09	0.59	
ADA questionnaire (21)	≥3	38	0.688	0.644-0.731	66	63	29	7.8	98	1.80	0.53	0.697
	≥5	5			17	95	12	14.7	96	3.68	0.87	
Brazilian score (35)	≥12	40	0.683	0.640-0.727	68	61	29	7.5	98	1.73	0.53	0.554
	≥18	15			30	86	16	9.1	96	2.12	0.81	
Oman score (39)	≥9	57	0.680	0.639-0.722	82	45	27	6.5	98	1.49	0.40	0.479
	≥10	39			64	62	26	7.3	97	1.67	0.58	
British score (14)	≥4	44	0.670	0.627-0.714	69	58	27	7.1	98	1.63	0.54	0.187
	≥6	16			32	85	17	9.0	96	2.10	0.80	
French score (DESIR study) (29)	≥3**	26	0.654	0.608-0.699	51	75	27	8.9	97	2.07	0.65	0.006
Rotterdam model (38)	≥32.9	47	0.646	0.599-0.693	71	54	25	6.8	98	1.55	0.53	0.035
	≥37	32			50	69	19	7.1	97	1.63	0.72	
Kuwait score (40)	≥22	40	0.630	0.587-0.674	59	61	20	6.6	97	1.50	0.68	0.002
	≥32	8			18	92	11	10.0	96	2.37	0.88	

Abbreviations: AROC, area under the curve of receiver operating characteristic curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; LR+, positive likelihood ratio; LR-, negative likelihood ratio.

Risk prediction models other than Korean score were arranged in the order of higher AROC value. The performance of different risk prediction models for the prediction of incident diabetes were evaluated with each model's original cut-off value and new cut-off value showing the highest Youden index. The results in the upper row of each score are based on the new cut-off value with the highest Youden index. The results in the bottom row of each score are based on the original cut-off value. *P value for the comparison of ROC curves between Korean score and other models. **The new cut-off and original cut-off were same for Japanese score and French score.

Figure 4. ROC curves of risk prediction models in longitudinal validation.



ROC curves of the Korean score, the 2 scores with the highest AROC, and the 2 scores with the lowest AROC.

Comparison between laboratory parameters and Korean screening score

We performed univariate and multivariate logistic regression analysis with the laboratory parameters (FPG, HbA1c, total cholesterol, HDL cholesterol, triglyceride) for the prediction of the incident diabetes (Table 5). All laboratory parameters showed significant association with incident diabetes in the univariate analysis. In the multivariate analysis with stepwise elimination of insignificant parameters (cut-off of P value >0.05), only FPG and HbA1c were significant predictors of incident diabetes. The estimated AROC of the univariate logistic model of FPG or HbA1c and the multivariate model of FPG and HbA1c was 0.771 (95% CI 0.729-0.813), 0.796 (95% CI 0.758-0.834), and 0.838 (95% CI 0.804-0.871), respectively. All these models had significantly higher AROC than the Korean score (Table 6, Figure 3).

To compare the risk classification for incident diabetes, we calculated the improvement of the risk classification by the models of laboratory parameters compare to the Korean score with the measures of NRI and IDI (Table 6). The NRIs based on the risk categories of <5%, ≤5 - <10%, ≤10 - <15%, and 15%≤ (24) were 27.3% (95% CI 13.9-40.6%) for FPG model, 27.8% (95% CI 14.3-41.3%) for HbA1c model, and 45.2% (95% CI 31.9-58.5%) for multivariate model of FPG and HbA1c. The addition of Korean score to the combined model of FPG and HbA1c increased the AROC (0.849, 95% CI 0.818-0.880) with small increment than the model without the Korean score, but did not improve the risk classification (NRI 4.6%, 95% CI -3.5-12.7%) (Table 6).

Table 5. Univariate and multivariate logistic regression analysis of laboratory parameters for incident diabetes.

Laboratory parameters	Univariate analysis		Multivariate analysis		
	Odds ratio (95% CI)	P value	β coefficient	Odds ratio (95% CI)	P value
FPG	1.11 (1.09-1.13)	<0.001	0.074	1.077 (1.056-1.097)	<0.001
HbA1c	37.406 (20.767-67.377)	<0.001	2.876	17.738 (9.408-33.443)	<0.001
Total cholesterol	1.008 (1.003-1.013)	<0.001			
HDL cholesterol	0.971 (0.957-0.985)	<0.001			
Triglyceride	1.005 (1.003-1.006)	<0.001			

Abbreviations: FPG, fasting plasma glucose; HbA1c, hemoglobin A1c

Above logistic regression analysis was applied to the study population for longitudinal validation (N = 3029).

Table 6. Comparison between Korean score and risk prediction models of laboratory parameters

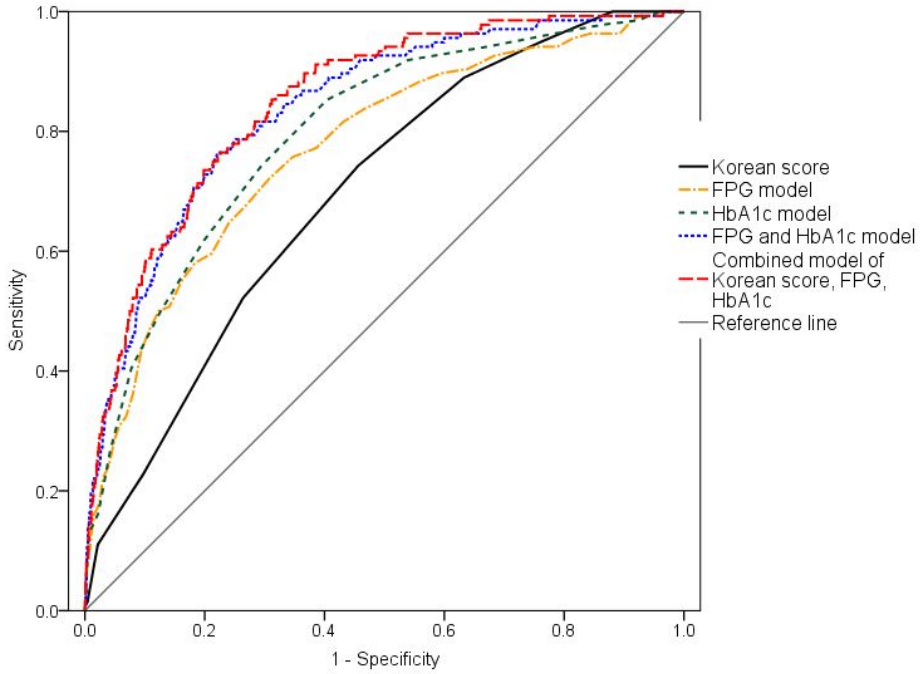
	Korean score	FPG model	HbA1c model	FPG and HbA1c model	Combined model of Korean score, FPG and HbA1c
AROC (95% CI)	0.696 (0.655-0.737)	0.771 (0.729-0.813)	0.796 (0.758-0.834)	0.838 (0.804-0.871)	0.849 (0.818-0.880)
Change in AROC* (P value)		0.075 (0.011)	0.100 (<0.001)	0.142 (<0.001)	0.153 (<0.001)
NRI (%)* (P value)		27.3 (<0.001)	27.8 (<0.001)	45.2 (<0.001)	52.1 (<0.001)
IDI* (P value)		0.046 (<0.001)	0.058 (<0.001)	0.099 (<0.001)	0.105 (<0.001)
Change in AROC** (P value)					0.011 (0.016)
NRI (%)** (P value)					4.6 (0.264)
IDI** (P value)					0.006 (0.176)

Abbreviations: AROC, area under the curve of receiver operating characteristic curve; FPG, fasting plasma glucose; HbA1c, hemoglobin A1c; NRI, net reclassification index; IDI, integrative discrimination improvement.

Above comparison analysis was done in the study population of longitudinal validation (N = 3029).

*Change in AROC, NRI and IDI were calculated for each model compared to Korean score. **Change in AROC, NRI and IDI were calculated to compare FPG and HbA1c model and combined model of Korean score, FPG and HbA1c.

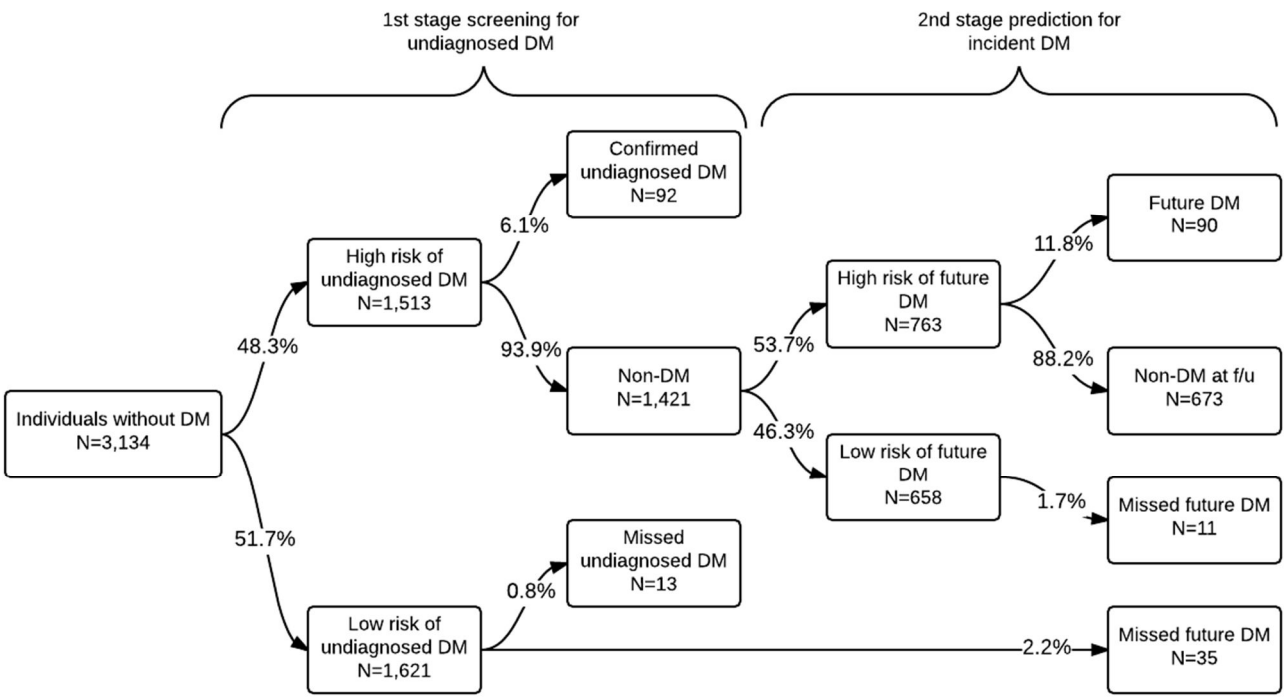
Figure 5. ROC curves of the Korean score and risk prediction models of laboratory parameters.



Simulation of diabetes screening

For the simulation of diabetes screening in a real clinical setting, first, we applied the Korean screening score to the population without known diabetes at baseline visit of longitudinal study. For the next step, the combined risk prediction model of the Korean score and laboratory parameters was applied to the high risk group classified by the Korean score. Among 3134 individuals without known diabetes at baseline, 1513 (48.3%) individuals were predicted as high risk of undiagnosed diabetes after application of the Korean score. Among the 1513 high risk individuals, the laboratory test of FPG and HbA1c confirmed diabetes in 92 (6.1%) individuals. Among the remaining 1421 individuals in the high risk group, the combined risk prediction model of Korean score, FPG and HbA1c classified 763 (53.7%) individuals as high risk of incident diabetes. Among the 763 individuals who were at high risk of incident diabetes, 90 (11.8%) individuals developed incident diabetes. The above result is summarized in the figure 6. In the first step screening, Korean score misclassified 13 (12.4%) undiagnosed diabetes as low risk and in the second step screening, the combined risk prediction model misclassified 11 (10.9%) incident diabetes as low risk.

Figure 6. Simulation of diabetes screening using risk prediction models



DISCUSSION

In this external validation of diabetes risk prediction models in a study population of health screening program participants, the Korean screening score composed of non-laboratory parameters demonstrated reasonable performance for the screening of undiagnosed diabetes, but limited ability to predict incident diabetes. In the comparison between the Korean screening score and laboratory parameters for the prediction of incident diabetes, the risk prediction model composed of FPG and HbA1c clearly showed higher discrimination index which was calculated as AROC, and improved risk classification than the Korean screening score. The addition of the Korean screening score to the risk prediction model of FPG and HbA1c increased the discrimination index, but only by small increment and with no improvement in risk classification.

Although the Korean screening score showed reasonable performance for screening undiagnosed diabetes compare to the other risk prediction models, it was not the prediction model with the highest discrimination index among the 17 models validated in this study. Some existing risk prediction models showed better performance than Korean score for undiagnosed and incident diabetes despite of different ethnicity and population characteristics in which they were developed. It was reported that risk prediction models developed in western countries poorly worked for Asian population (25), but in this study, some risk prediction models developed in western populations (19, 20, 26)

showed higher discrimination index than Korean score for undiagnosed and incident diabetes. This result suggests that same ethnicity or nationality does not guarantee the generalizability of a risk prediction model. External validation and recalibration of the model should be taken before its application to the different populations.

For screening of undiagnosed diabetes, the risk prediction models of non-laboratory parameters showed acceptable performance. Although, the positive predictive value was below 6%, the negative predictive value was as high as 99%. This suggests that the low risk group classified by the non-laboratory risk prediction model has very low possibility of having undiagnosed diabetes. However, the risk prediction model composed of basic laboratory parameters like FPG and HbA1c demonstrated superior performance for the prediction of incident diabetes than non-laboratory models. Additional blood test for FPG and HbA1c can be justified for more accurate prediction of future diabetes. However, testing FPG and HbA1c of a whole population as a mass-screening program needs considerable amount costs. A two-step approach combining non-laboratory screening and laboratory screening can be a solution for this problem (14). Because risk prediction models of non-laboratory parameters need virtually no cost, they can be applied to whole population to screen undiagnosed diabetes. Then, for those who have high risk of undiagnosed diabetes, simple blood test of FPG and HbA1c can detect or rule out true diabetes. If the test results do not meet the criteria of diabetes, the predicted risk of incident diabetes can be calculated and guide to determining who needs more intensive diabetes prevention program and more frequent follow-

up visits. This two-step approach can prevent unnecessary blood tests for low risk individuals and reduce the cost of screening program. Several previous studies also suggested a stepwise approach as an efficient diabetes screening program (14, 27). National Institute for Health and Care Excellence (NICE) guideline in UK recommends a two-stage screening (28). In detail, the first stage of screening is risk assessment using a validated risk-assessment tool or self-assessment questionnaire for adults with age over 40 years. For those with high risk of diabetes at first stage, either FPG or HbA1c should be tested. When the test results meet the diagnostic criteria of diabetes, the diabetes treatment should be initiated, and for those who have the test result of prediabetic range, the intensive lifestyle change should be offered (28). In our simulation of the two-step approach, 13/105 (12.4%) of undiagnosed diabetes and 11/101 (10.9%) of incident diabetes were misclassified as low risk. This percentage is not negligible, but further adjustment of the cut-off values would reduce the misclassification because the current cut-off values were only determined by Youden index without any clinical consideration. Addition to this, a comprehensive cost-effectiveness analysis should precede its application to the daily practice.

To further improve the prediction of incident diabetes, many biomarkers and genetic markers were analyzed. Each of these parameters was independent risk factor for incident diabetes, but these newer risk factors only modestly improved the risk prediction when they were added to the basic clinical parameters, such as family history, medical history, obesity and blood glucose level (24, 29-31). In this era of whole genome level study, more than dozens

of genetic markers have been discovered as risk factors of diabetes (32). However, the combination of various genetic markers was inferior to the basic non-laboratory parameters for the risk prediction and only marginally increased the discrimination index in addition to the basic clinical parameters (32). The evidence to date does not support the routine use of genetic markers for the prediction of diabetes (32), but with rapidly increasing new discoveries in the genetics and pathophysiology of diabetes, newer biomarkers or genetic markers can be a part of screening tool for diabetes.

One of the limitations of this study is that oral glucose tolerance test was not included in the definition of diabetes. Because FPG and HbA1c are less sensitive than oral glucose tolerance test (OGTT) in diagnosing diabetes (33), the prevalence of undiagnosed diabetes could be underestimated. However, because OGTT requires more time and cost than a fasting blood test, it is difficult to be used at a mass screening program. The second limitation is that the risk prediction models of laboratory parameters were not externally validated. The development and validation of the models were not done in independent cohorts. This could cause overestimation of the prediction performance, compare to the risk prediction models of non-laboratory parameters which were externally validated in our cohort. Another limitation is that the study population is not a good representative of general Korean population because participation of health screening program can be affected by individual's socioeconomic status and health-seeking behavior. The result of this study can be applied to other population in the setting of private health screening program in urban area, which is quite popular in Korea.

In conclusion, the risk prediction models composed of non-laboratory parameters including the Korean screening score can be a useful screening too for undiagnosed diabetes. However, their performance was inferior to the laboratory parameters for the prediction of incident diabetes. The measurement of fasting plasma glucose and hemoglobin A1c can significantly facilitate the risk prediction of incident diabetes.

REFERENCES

1. Guariguata L, Whiting D, Hambleton I, Beagley J, Linnenkamp U, Shaw J. Global estimates of diabetes prevalence for 2013 and projections for 2035. Diabetes research and clinical practice. 2014;103(2):137-49.

2. Patel A, MacMahon S, Chalmers J, Neal B, Billot L, Woodward M, et al. Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. N Engl J Med. 2008 Jun 12;358(24):2560-72. PubMed PMID: 18539916. Epub 2008/06/10. eng.

3. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. The Lancet. 1998 Sep 12;352(9131):837-53. PubMed PMID: 9742976. Epub 1998/09/22. eng.

4. Fodor JG, Adamo KB. Prevention of type 2 diabetes mellitus by changes in lifestyle. N Engl J Med. 2001;345(9):696.

5. Chiasson J-L, Josse RG, Gomis R, Hanefeld M, Karasik A, Laakso M. Acarbose for prevention of type 2 diabetes mellitus: the STOP-NIDDM randomised trial. The Lancet. 2002;359(9323):2072-7.

6. Gillies CL, Abrams KR, Lambert PC, Cooper NJ, Sutton AJ, Hsu RT, et al. Pharmacological and lifestyle interventions to prevent or delay type 2 diabetes in people with impaired glucose tolerance: systematic review and meta-analysis. British Medical Journal. 2007;334(7588):299.

7. Schellenberg ES, Dryden DM, Vandermeer B, Ha C, Korownyk C. Lifestyle Interventions for Patients With and at Risk for Type 2 DiabetesA Systematic Review and Meta-analysis. Annals of internal medicine. 2013;159(8):543-51.

8. Hemmingsen B, Lund SS, Gluud C, Vaag A, Almdal TP, Hemmingsen C, et al. Targeting intensive glycaemic control versus targeting conventional glycaemic control for type 2 diabetes mellitus. The Cochrane database of systematic reviews. 2013;11:CD008143. PubMed PMID: 24214280. Epub 2013/11/12. eng.

9. US Department of Health and Human Services. National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States. US Department of Health and Human Services. 2014.

10. Korean Ministry of Health and Welfare. Korea Health Statistics 2012: Korea National Health and Nutrition Examination Survey (KNHANESV-3). Seoul, South Korea, Ministry of Health and Welfare. 2012.

11. 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. British Medical Journal. 2012;345.

12. Collins GS, Mallett S, Omar O, Yu L-M. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. BMC medicine. 2011;9(1):103.

13. Noble D, Mathur R, Dent T, Meads C, Greenhalgh T. Risk models and scores for type 2 diabetes: systematic review. British Medical Journal. 2011;343.

14. Wannamethee S, Papacosta O, Whincup P, Thomas M, Carson C, Lawlor D, et al. The potential for a two-stage diabetes risk algorithm combining non-laboratory-based scores with subsequent routine non-fasting blood tests: results from prospective studies in older men and women. Diabetic Medicine. 2011;28(1):23-30.

15. Ninomiya T, Hata J, Hirakawa Y, Mukai N, Iwase M, Kiyohara Y. Two risk score models for predicting incident Type 2 diabetes in Japan. Diabetic Medicine. 2012;29(1):107-14.

16. 건강 iN homepage: National Health Insurance; 2013. Available from: hi.nhis.or.kr.

17. Lee Y-h, Bang H, Kim HC, Kim HM, Park SW, Kim DJ. A Simple Screening Score for Diabetes for the Korean Population Development, validation, and comparison with other scores. Diabetes Care. 2012;35(8):1723-30.

18. Association AD. Standards of medical care in diabetes 2010. Diabetes Care. 2010;33(Supplement 1):S11-S61.

19. 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. Medical Journal of Australia. 2010;192(4):197.

20. Glümer C, Carstensen B, Sandbæk A, Lauritzen T, Jørgensen T, Borch-Johnsen K. A Danish Diabetes Risk Score for Targeted Screening The Inter99 study. Diabetes Care. 2004;27(3):727-33.

21. Bang H, Edwards AM, Bomback AS, Ballantyne CM, Brillon D, Callahan MA, et al. Development and validation of a patient self-assessment score for diabetes risk. Annals of internal medicine. 2009;151(11):775-83.

22. Heianza Y, Arase Y, Saito K, Hsieh SD, Tsuji H, Kodama S, et al. Development of a screening score for undiagnosed diabetes and its application in estimating absolute risk of future type 2 diabetes in Japan: Toranomon Hospital Health Management Center Study 10 (TOPICS 10). The Journal of Clinical Endocrinology & Metabolism. 2013;98(3):1051-60.

23. Gray L, Taub N, Khunti K, Gardiner E, Hiles S, Webb D, et al. The Leicester Risk Assessment score for detecting undiagnosed type 2 diabetes and impaired glucose regulation for use in a multiethnic UK setting. Diabetic Medicine. 2010;27(8):887-95.

24. Talmud PJ, Hingorani AD, Cooper JA, Marmot MG, Brunner EJ, Kumari M, et al. Utility of genetic and non-genetic risk factors in prediction of type 2 diabetes: Whitehall II prospective cohort study. British Medical Journal. 2010;340.

25. Glümer C, Vistisen D, Borch-Johnsen K, Colagiuri S. Risk scores for type 2 diabetes can be applied in some populations but not all. Diabetes Care. 2006;29(2):410-4.

26. Alssema M, Vistisen D, Heymans M, Nijpels G, Glümer C, Zimmet P, et al. The Evaluation of Screening and Early Detection Strategies for Type

2 Diabetes and Impaired Glucose Tolerance (DETECT-2) update of the Finnish diabetes risk score for prediction of incident type 2 diabetes. Diabetologia. 2011;54(5):1004-12.

27. Echouffo-Tcheugui JB, Simmons RK, Williams KM, Barling RS, Prevost AT, Kinmonth AL, et al. The ADDITION-Cambridge trial protocol: a cluster–randomised controlled trial of screening for type 2 diabetes and intensive treatment for screen-detected patients. BMC Public Health. 2009;9(1):136.

28. Chatterton H, Younger T, Fischer A, Khunti K. Risk identification and interventions to prevent type 2 diabetes in adults at high risk: summary of NICE guidance. British Medical Journal. 2012;345:e4624. PubMed PMID: 22791792. Epub 2012/07/14. eng.

29. Balkau B, Lange C, Fezeu L, Tichet J, de Lauzon-Guillain B, Czernichow S, et al. Predicting Diabetes: Clinical, Biological, and Genetic Approaches Data from the Epidemiological Study on the Insulin Resistance Syndrome (DESIR). Diabetes Care. 2008;31(10):2056-61.

30. Raynor L, Pankow JS, Duncan BB, Schmidt MI, Hoogeveen RC, Pereira MA, et al. Novel risk factors and the prediction of type 2 diabetes in the Atherosclerosis Risk in Communities (ARIC) study. Diabetes Care. 2013;36(1):70-6.

31. Meigs JB, Shrader P, Sullivan LM, McAteer JB, Fox CS, Dupuis J, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. N Engl J Med. 2008;359(21):2208-19.

32. Vassy JL, Meigs JB. Is genetic testing useful to predict type 2 diabetes? Best practice & research Clinical endocrinology & metabolism. 2012 Apr;26(2):189-201. PubMed PMID: 22498248. Pubmed Central PMCID: Pmc4070012. Epub 2012/04/14. eng.

33. Phillips LS, Ratner RE, Buse JB, Kahn SE. We can change the natural history of type 2 diabetes. Diabetes Care. 2014 Oct;37(10):2668-76. PubMed PMID: 25249668. Pubmed Central PMCID: PMC4170125. Epub 2014/09/25. eng.

34. Aekplakorn W, Bunnag P, Woodward M, Sritara P, Cheepudomwit S, Yamwong S, et al. A risk score for predicting incident diabetes in the Thai population. Diabetes Care. 2006;29(8):1872-7.

35. de Sousa AGP, Pereira AC, Marquezine GF, do Nascimento-Neto RM, Freitas SN, Nicolato RLdC, et al. Derivation and external validation of a simple prediction model for the diagnosis of type 2 diabetes mellitus in the Brazilian urban population. European journal of epidemiology. 2009;24(2):101-9.

36. Chaturvedi V, Reddy K, Prabhakaran D, Jeemon P, Ramakrishnan L, Shah P, et al. Development of a clinical risk score in predicting undiagnosed diabetes in urban Asian Indian adults: a population-based study. CVD prevention and control. 2008;3(3):141-51.

37. Gao W, Dong Y, Pang Z, Nan H, Wang S, Ren J, et al. A simple Chinese risk score for undiagnosed diabetes. Diabetic Medicine. 2010;27(3):274-81.

38. Baan CA, Ruige JB, Stolk RP, Witteman J, Dekker JM, Heine RJ, et al. Performance of a predictive model to identify undiagnosed diabetes in a health care setting. Diabetes Care. 1999;22(2):213-9.

39. Al-Lawati J, Tuomilehto J. Diabetes risk score in Oman: a tool to identify prevalent type 2 diabetes among Arabs of the Middle East. Diabetes research and clinical practice. 2007;77(3):438-44.

40. Al Khalaf MM, Eid MM, Najjar HA, Alhajry KM, Doi SA, Thalib L. Screening for diabetes in Kuwait and evaluation of risk scores. East Mediterrian Health Journal. 2010;16(7):725-31.

국문 초록

서론: 진단되지 않은 당뇨를 조기 진단하고 미래의 당뇨병 발생을 예측하는 것은 당뇨병으로 인한 심각한 후유증을 예방하고 늦추기 위해서 매우 중요하다. 이를 위해 다양한 당뇨 위험 예측 모델이 개발되었고 최근에는 한국인을 대상으로 한 당뇨병 위험 자가 측정 모델이 만들어졌다. 그러나 한국인을 대상으로 한 당뇨병 위험 예측 모델은 진단되지 않은 당뇨병 환자를 예측하는 목적으로만 검증되었고 미래의 당뇨병 발생을 예측하는 목적으로는 아직 검증된 바가 없다. 이에 본 연구에서는 한국인 당뇨병 위험 예측 모델이 미래의 당뇨병 발생을 예측하는 예측력을 검증하고 이를 다른 당뇨병 위험 예측 모델과 실험실 검사 결과의 예측력과 비교하고자 한다.

방법: 2011 년 1 월부터 2012 년 9 월까지 서울대학교병원 강남센터에 내원한 26,675 명의 진료 기록을 진단되지 않은 당뇨 환자를 선별하는 단면적 연구에 활용하였고 2004 년 1 월부터 2005 년 12 월까지 서울대학교병원 강남센터를 내원하고 2011 년 1 월부터 2012 년 9 월 사이에 재내원한 기록이 있는 3,029 명의 진료기록을 미래의 당뇨병 발생을 분석하는 종적 연구에 활용하였다. 한국인 당뇨병 위험 예측 모델과 지금까지 출판된 16 개의 실험실 검사 결과를 포함하지 않는 다른 위험 예측 모델을 분석하였고 실험실 검사 결과

를 바탕으로 새로운 당뇨병 위험 예측 모델을 만들어 이들의 예측력도 비교 분석하였다.

결과: 진단되지 않은 당뇨병을 선별하는 단면적 연구에서 한국인 당뇨병 위험 예측 모델의 민감도는 81%, 특이도는 58%, ROC 곡선 아래의 면적은 0.754 로 나타났다. 다른 위험 예측 모델도 비슷한 정도의 ROC 곡선 아래의 면적 값을 나타냈다. 미래의 당뇨병 발생을 예측하는 종적 연구에서 한국인 모델의 민감도는 74%, 특이도는 54%, ROC 곡선 아래 면적은 0.696 으로 나타났다. 이와 비교하여, 공복 혈당과 헤모글로빈 A1c 수치로 구성된 당뇨병 위험 예측 모델은 유의하게 높은 ROC 곡선 아래 면적 값 (0.838 vs. 0.696, P value <0.001) 을 나타냈다. 공복혈당과 헤모글로빈 A1 로 이루어진 예측 모델에 한국인 당뇨병 예측 모델 점수를 결합하면 ROC 곡선 아래 면적 값은 적은 정도로 상승하나 (0.838 vs. 0.849, P value = 0.016) 위험도 재분류에 있어서는 유의한 개선이 없었다 (net reclassification index 4.6%, P value = 0.264).

결론: 결론적으로 한국인 당뇨병 위험 예측 모델은 진단되지 않은 당뇨 환자를 선별하는 데에는 유용하게 사용될 수 있으나 미래의 당뇨병 발생을 예측하는 데에는 실험실 검사 결과를 활용하는 것보다 예측력이 열등하다.

주요어 : 당뇨병, 위험 예측, 검증 연구

학 번 : 2013-21733