The official journal of the Japan Atherosclerosis Society and
the Asian Pacific Society of Atherosclerosis and Vascular Diseases

Original Article

J Atheroscler Thromb, 2020; 27: 1160-1175. http://doi.org/10.5551/jat.48843

Development of a Cardiovascular Disease Risk Prediction Model Using the Suita Study, a Population-Based Prospective Cohort Study in Japan

Michikazu Nakai[1], Makoto Watanabe[2], Yoshihiro Kokubo[2], Kunihiro Nishimura[3], Aya Higashiyama[2, 3], Misa Takegami[3], Yoko M Nakao[2, 3], Tomonori Okamura[4] and Yoshihiro Miyamoto[1, 2]

[1]Center for Cerebral and Cardiovascular Disease Information, National Cerebral and Cardiovascular Center, Suita, Japan
[2]Department of Preventive Cardiology, National Cerebral and Cardiovascular Center, Suita, Japan
[3]Department of Preventive Medicine and Epidemiologic Informatics, National Cerebral and Cardiovascular Center, Suita, Japan
[4]Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan

Aim: To construct a risk prediction model for cardiovascular disease (CVD) based on the Suita study, an urban Japanese cohort study, and compare its accuracy against the Framingham CVD risk score (FRS) model.

Methods: After excluding participants with missing data or those who lost to follow-up, this study consisted of 3,080 men and 3,470 women participants aged 30–79 years without CVD at baseline in 1989–1999. The main outcome of this study was incidence of CVD, defined as the incidence of stroke or coronary heart disease. Multivariable Cox proportional hazards models with stepwise selection were used to develop the prediction model. To assess model performance, concordance statistics (C-statistics) and their 95% confidence intervals (CIs) were calculated using a bootstrap procedure. A calibration test was also conducted.

Results: During a median follow-up period of 16.9 years, 351 men and 241 women developed CVD. We formulated risk models with and without electrocardiogram (ECG) data that included age, sex, systolic blood pressure, diastolic blood pressure, high-density lipoprotein cholesterol, non-high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, diabetes mellitus, smoking, and urinary protein as risk factors. The C-statistics of the Suita CVD risk models with ECG data (0.782; 95% CI, 0.766–0.799) and without ECG data (0.781; 95% CI, 0.765–0.797) were significantly higher than that of the FRS model (0.768; 95% CI, 0.750–0.785).

Conclusions: The Suita CVD risk model is feasible to use and improves predictability of the incidence of CVD relative to the FRS model in Japan.

Key words: Cohort studies, Risk score model, Stroke, Coronary heart disease

Introduction

Cardiovascular diseases (CVDs) continue to be the leading causes of mortality worldwide[1, 2]. CVD prevention is one of the most urgent concerns to be addressed. One aspect of CVD prevention is to identify individuals with higher risk for CVDs and manage their risk factors appropriately. Accordingly, many risk prediction tools for CVD components have been developed, such as the Framingham CVD risk score (FRS) model[3] and the Suita score model for coronary heart diseases (CHDs)[4]. However, it is usually not convenient for physicians to calculate risks separately by component and integrate them.

Recently, several risk prediction tools for multiple CVDs (CHDs and strokes) that are simple to use have been developed. However, such risk prediction tools for multiple CVDs depend on the proportion of CVD components. In other words, most of these risk prediction tools are based mainly on data from West-

Address for correspondence: Michikazu Nakai, Center for Cerebral and Cardiovascular Disease Information, National Cerebral and Cardiovascular Center, 6-1 shinmachi kishibe, Suita-city, Osaka, JAPAN 564-8565 E-mail: nakai.michikazu@ncvc.go.jp

Received: January 23, 2019 Accepted for publication: December 19, 2019

ern countries[3, 5-9], where CHDs develop more frequently than strokes. In Asians, strokes develop more commonly than CHDs[10]. In particular, Japan is a country with a remarkably lower incidence of CHDs[11]. Japanese cohort studies show that strokes develop two times more frequently than CHDs. Stroke prevention remains an essential concern in Japan[10, 11].

In the past, two risk prediction tools for multiple CVDs have been developed in Japan, one is based on the Hisayama study[12, 13] of CVD incidence in rural southwestern Japan and the other is on the NIPPON-DATA80 study[14] of CVD mortality in a representative Japanese population. However, the Hisayama risk prediction tool with a baseline survey from 1988 may not be applicable to the current Japanese population, as most are living in urban areas and rapidly acquiring westernized lifestyles[12, 13, 15, 16]. Similarly, it would be difficult to use the NIPPONDATA80 risk prediction tool nowadays. Its baseline survey occurred in 1980, and its risk model adopted CVD mortality as the outcome. Age-adjusted CVD mortality has declined considerably in the past several decades[11, 17]. In addition, disease mortality is influenced by advances in therapeutic technologies and agents as well as disease incidence. For greater applicability to more populations, it would be important to make a risk prediction tool with disease incidence reflective of the current urban Japanese population.

Consequently, we have developed a 10-year risk prediction model for multiple CVDs in the Suita study, a recent urban Japanese cohort study. We have compared its accuracy against the FRS model[3]. Moreover, we compared risk models with and without data from electrocardiograms (ECGs), which are often included in health checkups by law in Japan. ECGs can provide useful information for early detection of CVD risk factors such as atrial fibrillation (AF)[18].

Methods

Populations

The Suita study is a cohort study about CVD among urban residents in Suita, Japan, that began in 1989. The details of this study have been described elsewhere[19-24]. Briefly, 12,200 and 3,000 participants aged 30 to 79 years were randomly selected in 1989 and 1996, respectively, from the municipality population registry of Suita City. Participants were stratified by sex and 10-year age categories. There were 7,808 participants (original cohort: 6,480; secondary cohort: 1,328) who received baseline regular health checkups between September 1989 and March 1996 (original cohort) or April 1996 and March 1999 (secondary cohort) at the National Cardiovascular Center (NCVC). They have returned every two years for follow-up examinations. Informed consent was given by all participants.

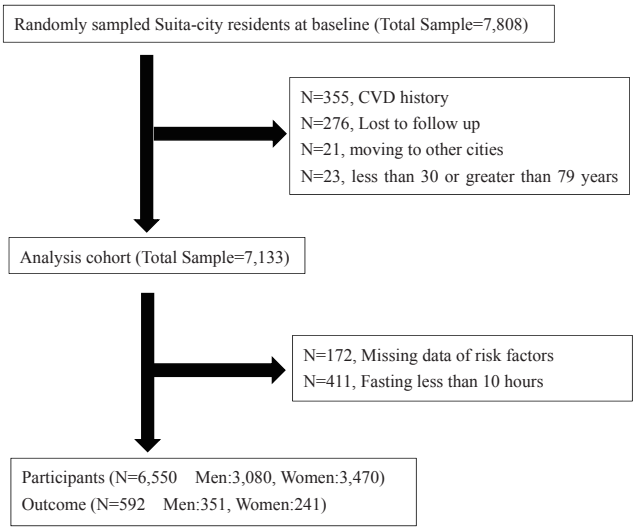


Fig. 1. Flow-chart of study participants

We excluded 1,258 participants because they had a previous history of CVD (n=355), were lost to follow-up (n=276), had moved to other cities before the baseline examination (n=21), were younger than 30 or older than 79 years at baseline (n=23), had missing data about risk factors (n=172), or did not fast for ≥ 10 hours before blood test (n=411). The remaining 6,550 participants (3,080 men and 3,470 women) were included in the analysis (**Fig. 1**). This cohort study was approved by the NCVC institutional review board in Suita, Japan (M17-001).

Baseline Examinations

Blood samples were centrifuged immediately after collection. Routine blood examinations were performed, including measurement of total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), and triglycerides (TG). Well-trained physicians measured blood pressure three times after 5 minutes of rest using a standard mercury sphygmomanometer. The average of the second and third measurements was used for the analysis. Public health nurses obtained information on cigarette smoking status, alcohol drinking habits, and medical history. The categorization of HDL-C in this study was based on criteria in the Japan Atherosclerosis Society (JAS) guidelines[25].

In addition to classifying blood pressure according to the Japanese Society of Hypertension classification system[26], we verified whether medication was

significantly effective. In **Supplemental Table 1**, we compared the following groups: (i) participants who took anti-hypertensive drugs and had decreased blood pressure, considered normal (systolic blood pressure (SBP) <140 mmHg and diastolic blood pressure (DBP) <90 mmHg), and (ii) participants who took anti-hypertensive drugs with no change in blood pressure, considered to have high blood pressure (SBP ≥ 140 mmHg or DBP ≥ 90 mmHg). There were no significant differences in both univariable ($P=0.554$) and multivariable ($P=0.712$ with ECG; $P=0.669$ without ECG) models. In addition, we compared (i) participants who took anti-hypertensive drugs and had decreased blood pressure, considered normal, and (iii) patients with Stage II hypertension (SBP ≥ 160 mmHg or DBP ≥ 90 mmHg). There were no statistical differences in univariable ($P=0.477$) and multivariable ($P=0.552$ with ECG; $P=0.551$ without ECG) models. Therefore, we have combined these groups into the same category if SBP ≥ 160 mmHg or DBP ≥ 100 mmHg or if they were taking medication.

Since CVD risk increases with age, categories for age were based on 10-year intervals up to age 60 years and 5-year intervals thereafter. Diabetes mellitus (DM) is defined as having a fasting blood glucose level of ≥ 126 mg/dL, currently using anti-diabetic medication, or both. Cigarette smoking was dichotomized as current versus other. Low-density lipoprotein cholesterol (LDL-C) was calculated using the Friedewald equation[27]. Non-HDL-C was calculated by subtracting HDL-C from TG. If the TG level was greater than 400 mg/dL, we set the value as missing ($n=96$).

A standard 12-lead ECG was obtained from all participants. Each ECG was coded independently using the Minnesota Code by two well-trained physicians. On the basis of the Minnesota Code classification system, AF was defined as 8-3-1, and left ventricular hypertrophy (LVH) was defined as both 3-1 and ST-T change (4-1-1, 4-1-2, 4-2, 4-3, 5-1, 5-2, or 5-3).

Definition of CVDs

The end of the follow-up period was on (1) the date of the first CHD or stroke event, (2) the date of death, (3) the date of leaving Suita City, or (4) December 31, 2013 (censored), whichever came first.

Follow-Up Survey

The health status of each participant was checked during clinical visits every 2 years and through a yearly questionnaire by mail or telephone. Medical records of participants with suspected CHD or stroke were reviewed by registered hospital physicians who were blinded to the baseline information. We defined fatal CHD (International Classification of Diseases 10th Revision (ICD-10) codes: I200–I259) and fatal stroke (ICD-10 code: I60–I699) on the basis of the underlying cause of death on death certificates provided by the Ministry of Health, Welfare, and Labour (MHWL). In Japan, the MHWL accumulates death certificates by law and codes them on the basis of ICD codes to compile national vital statistics. The criteria for myocardial infarction (MI) were based on the World Health Organization Monitoring of Trends and Determinants in Cardiovascular Disease project (MONICA)[28]. CHD consisted of coronary angioplasty, coronary artery bypass grafting, and sudden cardiac death within 24 hours. Stroke was defined according to criteria from the U.S. National Survey of Stroke[29], which was based on computed tomography findings. Strokes were classified on the basis of computed tomography, magnetic resonance imaging, and autopsy findings.

Statistical Analysis

Analysis of variance and χ^2 tests were used to compare means and frequencies. Age-adjusted multivariable Cox proportional hazards modeling was performed. Stepwise selection was used to develop the best risk prediction model. We constructed two models, one with ECG data and another without ECG data. Risk factors included sex, age, SBP, DBP, HDL-C, non-HDL-C, LDL-C, smoking status, DM, and urinary protein. When each value of blood pressures (SBP and DBP) or lipid profiles (LDL-C and non-HDL-C) belonged to different categories, we adopted the higher category to estimate hazard ratios, regression coefficients, or scores. To derive the risk score for each variable, coefficients were multiplied by a factor of 10 to allow for integer scores. The total score was calculated as the sum of the individual weighted scores. Total FRS was calculated according to CVD points for both men and women, as previously described[30]. The CVD risk prediction models were translated into a risk score sheet using methods developed in the Framingham Heart Study[30]. In order to estimate the 10-year risk probability for a CVD incident, the equation $1-S_0(t)^{\exp(x_i b)}$ is used, where $S_0(t)$ is the baseline survival function at the follow-up time ($t=10$ years) and $x_i b$ is a linear predictor of the total score for each participant from the fitted model. Then, the estimated cardiovascular risk score at 10 years for each category was calculated as the average 10-year risk for CVD events in the corresponding category.

Furthermore, we evaluated the discriminatory ability of the risk models using concordance statistics (C-statistics)[31-33] and 95% confidence intervals (CIs) with 200 iterations of a bootstrap method. Parameters were defined on the basis of the assumption that a

Table 1. Baseline characteristics and incident cardiovascular disease during the follow-up by sex

	Men (n = 3,080)	Women (n = 3,470)	P-value
Age, years	56 ± 13	54 ± 13	< 0.001
Body mass index kg/m^2	22.8 ± 2.9	22.2 ± 3.3	< 0.001
Systolic blood pressure, mmHg	128 ± 21	125 ± 22	< 0.001
Diastolic blood pressure, mmHg	80 ± 12	76 ± 12	< 0.001
Low density lipoprotein cholesterol, mg/dL	123 ± 33	133 ± 35	< 0.001
High density lipoprotein cholesterol, mg/dL	50 ± 14	58 ± 14	< 0.001
Non- high-density lipoprotein cholesterol, mg/dL	151 ± 35	154 ± 39	< 0.001
Creatinine, mg/dL	0.90 ± 0.28	0.68 ± 0.21	< 0.001
estimated glomerular filtration rate, mL/min/1.73 m^2	89.0 ± 32.1	92.9 ± 28.6	0.0061
Cigarette smoking, %	50.3%	11.9%	< 0.001
Count of cigarettes per day	24 ± 12	15 ± 9	< 0.001
Alcohol drinking, %	75.2%	32.4%	< 0.001
Antihypertensive drug, %	11.4%	10.7%	0.34
Diabetes mellitus, %	6.6%	3.6%	< 0.001
Urinary protein 1 + or more, %	20.6%	16.6%	< 0.001
Atrial Fibrillation, %	1.0%	0.3%	< 0.001
Left Ventricular Hypertrophy, %	1.9%	1.8%	0.77
Cumulative incidence of cardiovascular disease, %	11.4%	6.9%	< 0.001
Cumulative incidence of Stroke, %	6.4%	4.5%	0.001
Cumulative incidence of coronary heart disease, %	5.6%	2.6%	< 0.001

Means with standard deviations were shown

population of bivariate data pairs of the form (X_i, Y_i) exists. On the basis of a scheme for sampling pairs of pairs [(X_i, Y_i), (X_j, Y_j)] from that population, a pair of pairs was considered concordant if the larger X value was paired with the smaller Y value. The C-statistic represents the conditional probability of concordance plus half of the conditional probability that the data pairs are neither concordant nor discordant, assuming that the two Y-values can be ordered[33]. STATA commands somersd and lincom were used to calculate C-statistics and confidence limits and to compare the accuracy of our models and that of the FRS model on the basis of P values. To assess the calibration of our risk prediction models, we used the modified Hosmer–Lemeshow test (also called the Gronnesby and Borgan test) to evaluate predicted event percentage and actual observed event percentage within 10 years by deciles or sextiles of risk scores as a sub-analysis.

To explore the consistency of the risk score developed, 70% of the study participants (n=4,585) were randomly assigned to the risk prediction model as a derivation cohort, and the remaining 30% (n=1,965) were reserved to validate the risk score assigned by the derivation cohort. Finally, we evaluated a category-free version of net reclassification improvement (NRI (> 0)) proposed by Pencina *et al.*[34, 35] We followed the procedures as outlined by Goldman *et al.*[36].

All analyses were performed using the SAS software program, version 9.4 (SAS Institute, Cary, NC, USA), and the STATA software program, version 14.2 (StataCorp LP, College Station, TX, USA).

Results

During the median follow-up period of 16.9 years with 99,673 person-years (original cohort: median, 18.8 years with 84,870 person-years; secondary cohort: median, 16.1 years with 14,803 person-years), 351 men developed CVDs, consisting of 164 CHDs and 187 strokes. In addition, 241 women developed CVDs, consisting of 87 CHDs and 154 strokes. CHDs included 75 definite MIs, 22 possible MIs, 93 coronary interventions, 4 sudden cardiac deaths, and 57 fatal MIs on the basis of death certificate data only. Strokes consisted of 206 cerebral infarctions, 67 intracerebral hemorrhages, 30 subarachnoid hemorrhages, and 38 strokes of unknown type. Among strokes, 59 were detected on the basis of only death certificate data.

Table 1 shows risk factors by sex. All factors are significantly different between males and females except for anti-hypertensive medication and LVH. Age-adjusted multivariable hazard ratios from the risk prediction models are shown in Table 2. Incidence rates for each categorical variable are shown in Supplemental Table 2. Tables 3 and 4 provide risk score

Table 2. Hazard ratios for the cardiovascular risk prediction model with/ without Electrocardiogram

	Univariable with age adjustment	Interaction P with sex	Model with ECG	Model without ECG
			Multivariable	Multivariable
	HR (95% CI)		HR (95% CI)	HR (95% CI)
Men	1.76 (1.49, 2.08)		1.46 (1.21, 1.75)	1.46 (1.22, 1.76)
Age, years				
30-39			1.00	1.00
40-49		0.418	2.26 (1.27, 4.03)	2.27 (1.27, 4.04)
50-59		0.148	4.00 (2.32, 6.89)	4.05 (2.35, 6.98)
60-64		0.082	5.83 (3.36, 10.12)	5.96 (3.43, 10.34)
65-69		0.092	9.12 (5.23, 15.90)	9.27 (5.32, 16.16)
70-74		0.051	12.55 (7.22, 21.80)	13.10 (7.54, 22.74)
75-79		0.069	16.80 (9.45, 29.89)	17.36 (9.77, 30.86)
SBP < 120 mmHg and DBP < 80 mmHg	0.66 (0.52, 0.85)	0.933	0.68 (0.53, 0.88)	0.69 (0.54, 0.89)
SBP = 120-139 mmHg or DBP = 80-89 mmHg	1.00		1.00	1.00
SBP = 140-159 mmHg or DBP = 90-99 mmHg	1.34 (1.05, 1.69)	0.968	1.34 (1.05, 1.70)	1.34 (1.06, 1.70)
SBP ≥ 160 mmHg or DBP ≥ 100 mmHg or in-medication	1.84 (1.49, 2.25)	0.253	1.81 (1.47, 2.23)	1.84 (1.50, 2.27)
Non-HDL-C < 170 mg/dL and LDL-C < 140 mg/dL	1.00		1.00	1.00
Non-HDL-C ≥ 170 mg/dL or LDL-C ≥ 140 mg/dL	1.13 (0.96, 1.33)	0.918	1.19 (1.01, 1.41)	1.18 (1.00, 1.39)
HDL-C < 40 mg/dL	1.00		1.00	1.00
HDL-C = 40-59 mg/dL	0.71 (0.58, 0.88)	0.146	0.79 (0.64, 0.97)	0.79 (0.65, 0.98)
HDL-C ≥ 60 mg/dL	0.51 (0.40, 0.65)	0.074	0.66 (0.51, 0.85)	0.66 (0.51, 0.85)
Smoking	1.66 (1.39, 1.97)	0.258	1.43 (1.18, 1.74)	1.42 (1.17, 1.72)
Diabetes mellitus	2.36 (1.83, 3.05)	0.168	1.89 (1.45, 2.46)	1.95 (1.50, 2.53)
Urinary protein 1 + or more	1.54 (1.28, 1.86)	0.099	1.23 (1.02, 1.49)	1.25 (1.04, 1.52)
Atrial fibrillation	2.45 (1.31, 4.59)	0.257	2.29 (1.21, 4.33)	
Left ventricular hypertrophy	1.94 (1.32, 2.84)	0.179	1.64 (1.11, 2.42)	

Abbreviations: SBP, Systolic blood pressure; DBP, Diastolic blood pressure; LDL, Low density lipoprotein; HDL, high-density lipoprotein; Non-HDL, Non- high-density lipoprotein

When each value of blood pressures (SBP and DBP) or lipid profiles (LDL-C and non-HDL-C) belonged to different categories, we adopted the higher category to estimate hazard ratios.

sheets that can be used for predicting the 10-year risk for CVD.

The C-statistics for models with versus without ECG data were 0.782 (95% CI, 0.766–0.799) and 0.781 (95% CI, 0.765–0.797), respectively. The C-statistic for the FRS model was 0.768 (95% CI, 0.750–0.785). The Suita CVD risk models had statistically higher prediction accuracy than the FRS model ($P<0.001$ for both with and without ECG), which was applied to the Suita cohort.

Fig. 2 shows the calibration histograms comparing predicted event percentage and actual observed event percentage with and without ECG data. The χ^2-statistic comparing the risk prediction models with versus without ECG data was not significant, which indicated excellent goodness of fit. Furthermore, subgroup analysis by gender and age stratification for calibration histograms with and without ECG data was conducted (**Supplemental Fig. 1**). Calibration histograms were plotted in sextiles instead of deciles because of the sample size. Although the C-statistic for elderly participants (≥ 65 years old) was lower than that for other age categories (**Table 5**), the χ^2-statistic was not significant for models with versus without ECG data, which indicated acceptable goodness of fit. There were no statistically significant differences in prediction accuracy for models with versus without ECG data ($P=0.271$). The NRI of the two models was 0.06 (95% CI, −0.02 to 0.15). The P-value was 0.155, which indicated no statistical difference.

Supplemental Table 3 shows coefficients and risk scores with the derivation cohort ($n=4,585$) for the two models. The results were similar to those in **Table 3**. The C-statistic for models with and without ECG data in the validation cohort was 0.779 (95% CI, 0.760–0.798) and 0.778 (95% CI, 0.759–0.796),

Table 3. Multivariable regression coefficient and cardiovascular disease risk points with/ without Electrocardiogram

		Model with Electrocardiogram		Model without Electrocardiogram	
		Coefficient	Score	Coefficient	Score
Men		0.38	4	0.38	4
Age	30-39 years	0	0	0	0
	40-49 years	0.82	8	0.82	8
	50-59 years	1.39	14	1.40	14
	60-64 years	1.76	18	1.78	18
	65-69 years	2.21	22	2.23	22
	70-74 years	2.53	25	2.57	26
	75-79 years	2.82	28	2.85	29
BP	SBP < 120 mmHg and DBP < 80 mmHg	−0.38	−4	−0.37	−4
	SBP = 120-139 mmHg or DBP = 80-89 mmHg	0	0	0	0
	SBP = 140-159 mmHg or DBP = 90-99 mmHg	0.29	3	0.29	3
	SBP ≥ 160 mmHg or DBP ≥ 100 mmHg or in medication	0.59	6	0.61	6
Non-HDL-C < 170 mg/dL and LDL-C < 140 mg/dL		0	0	0	0
Non-HDL-C ≥ 170 mg/dL or LDL-C ≥ 140 mg/dL		0.17	2	0.17	2
HDL-C	HDL-C < 40 mg/dL	0	0	0	0
	HDL-C = 40-59 mg/dL	−0.23	−2	−0.23	−2
	HDL-C ≥ 60 mg/dL	−0.42	−4	−0.42	−4
Smoking		0.36	4	0.35	4
DM		0.64	6	0.67	7
Urinary protein 1 + or more		0.21	2	0.23	2
Atrial fibrillation		0.83	8		
Left ventricular hypertrophy		0.49	5		

When each value of blood pressures (SBP and DBP) or lipid profiles (LDL-C and non-HDL-C) belonged to different categories, we adopted the higher category to estimate regression coefficients and scores.

Table 4. Estimation in 10-year risks for cardiovascular diseases

Score	Probability (With ECG) baseline survival rate at 10 years = 0.99411	Probability (Without ECG) baseline survival rate at 10 years = 0.99389
0 or less	Less than 1%	Less than 1%
Score = 1-20	2%	2%
Score = 21-25	6%	6%
Score = 26-30	9%	9%
Score = 31-35	15%	14%
36 or more	26%	25%

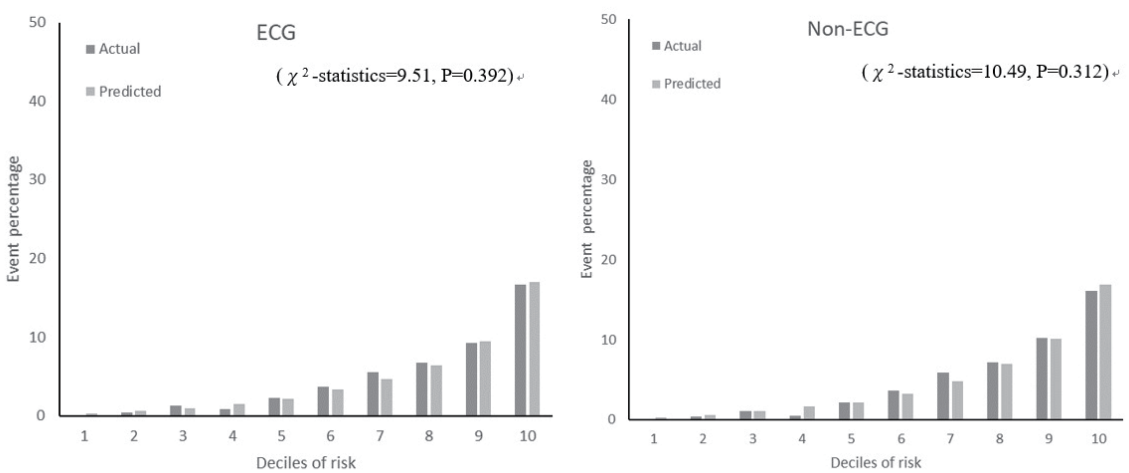


Fig. 2. Actual and predicted cardiovascular event percentage by deciles of risk with ECG and non-ECG

Table 5. C-statistics by sex and age categories (<65, ≥ 65) in the Suita CVD risk model with and without Electrocardiogram

	C-statistic (95% Confidence Interval)	
	Model with ECG	Model without ECG
Men	0.754 (0.732, 0.775)	0.752 (0.731, 0.774)
Women	0.794 (0.771, 0.818)	0.794 (0.770, 0.817)
Age ≥ 65 years old	0.645 (0.613, 0.678)	0.643 (0.610, 0.675)
Age <65 years old	0.764 (0.737, 0.790)	0.763 (0.737, 0.789)

respectively. Since the C-statistic of the FRS model that was applied to the Suita cohort was 0.765 (95% CI, 0.746–0.784), our models had higher accuracy (P =0.005 with ECG; P=0.012 without ECG). The calibration histograms comparing predicted event percentage and actual observed event percentage with versus without ECG data in the validation cohort were not significantly different (**Supplemental Fig. 2**). The models with versus without ECG data had similar prediction accuracy (P=0.436). The χ^2 test showed no significant differences between the model with ECG data and the model without ECG data (**Supplemental Fig. 3**).

Discussion

In this study, we developed a new 10-year Suita CVD risk model and demonstrated it had higher accuracy than the FRS model among urban Japanese individuals, who have a relatively high incidence of CHD but still lower than those in the United States. Since most of the Japanese population currently lives in urban areas[16], the Suita CVD risk model could be applicable to considerably more Japanese individuals, making it a useful tool for the prevention of CVDs.

According to the 2017 JAS guidelines, LDL-C and HDL-C levels should be tested during the initial screening in clinical settings. If the LDL-C value is above 120 mg/dL, non-HDL-C is carefully tested. Several papers in Japan also reported the relationship between the non-HDL-C value and incident CHDs[20, 21, 37]. Instead of using only the TC value, this Suita CVD risk model can select either non-HDL-C or LDL-C as factors. Therefore, the Suita CVD risk model would be very convenient because two risk factors for dyslipidemia can be applied simultaneously. Furthermore, this model is the first reported risk prediction tool with non-HDL-C in Japan, which could lead to increased use of non-HDL-C in primary care settings.

In addition, we constructed two risk models, a model with ECG data and a model without ECG data. According to C-statistics and NRI, there were no significant differences between the two models. ECG is often examined as part of health checkups in Japan. The FRS model also includes AF and LVH as important indexes[38]. Thus, it is important to determine if ECG should be included in the Japanese CVD risk model; this was the first attempt to use a Japanese cohort study. The results showed that the model without ECG data performed as accurately as the model with ECG in a Japanese CVD risk model. Applying the model with or without ECG data could depend on user preference.

The Suita CVD risk model was more applicable to Japanese individuals than the FRS model. Needless to say, this would be because of clinical differences between different populations, such as cholesterol levels or disease composition (stroke or CHD). The Hisayama study, a cohort study in rural Japan that began in the 1990s, presented a new CVD risk model with stroke and CHD[39]. The Suita CVD risk model included similar variables and had results similar to those of the Hisayama study. However, we are not able to directly compare these two models using statistical methods. Another report based on the Hisayama study[40] showed that the recent ratio of stroke to CHD incidence is decreasing (men 1.3, women 2.5), which was close to that in the Suita study (men 1.1, women 1.8). This indicates that differences in disease composition between rural and urban areas in Japan are getting smaller because lifestyles throughout Japan are becoming more westernized and homogenized.

Several limitations of this study should be discussed. First, this Suita CVD risk model was formulated on the basis of limited samples from a limited area. External validation with other cohort studies was not conducted. However, we did evaluate internal validation by generating histograms between predicted event percentage and actual observed event percentage and analyzed C-statistics using a bootstrap method. Second, the C-statistic in elderly participants was lower because of the small sample size. The proportion of elderly participants in this population was only 25.8% (n=1,689), which may have decreased power in the analysis. The goodness of fit test indicated a

good fit among elderly participants. Furthermore, Vliegenthart *et al.* have shown that coronary calcification improves CVD prediction among elderly individuals, but our cohort study did not collect data about coronary calcification[41]. Third, the generalizability of this risk score is limited. We cannot conclude that this risk model is applicable to all urban Japanese residents because the cohort study was based in a single city. However, as stated above, there is a possibility that the risk model in this study may be partly applicable to residents in rural areas since most of Japan has been westernizing rapidly in the past 30 years. Fourth, there were 3,271 participants who were censored for deaths from other causes (n=181), left Suita City (n=762), and withdrew from the study and with other reasons (n=2,328). Among participants who died, we might have missed CVD events before death if the participant did not answer yearly questionnaires or attended clinical visits for CVD follow-up. We might not have been able to detect CVD events after participants left Suita City or withdrew from the study. Consequently, such a large number of censored participants might have led to a decrease in the absolute risk for CVDs, thus unexpectedly biasing the results. Fifth, approximately 20% of CVD events were detected on the basis of the data on the underlying causes of death found only in death certificates. This might bias the results because of relatively low accuracy of diagnosis on death certificates. Additionally, the methods to fill in death certificates substantially changed in 1995, but the proportions of CVD events detected with death certificates only were similar before and after 1995. We think such major changes did not systematically impact the analysis results. Finally, the Suita CVD risk model did not include chronic kidney disease, which was included in a previous Suita CHD risk model[4]. In this study, since the baseline period was extended to 1999, there were two methods used to measure serum creatinine. It is possible to calculate estimated glomerular filtration rate with each method using different equations, but the creatinine measurement methods were not compatible and results could not be standardized.

Conclusion

We developed a 10-year Suita CVD risk model for an urban Japanese population. Compared with a previous model, ours has better predictability and feasibility. This new risk score model could be widely applicable to urbanized Japanese individuals. It is feasible enough to be used during health checkups or in primary care settings in Japan. Further studies with relatively large samples should be conducted for external validation of this model.

Conflict of Interest

Kunihiro Nishimura received research findings from Philips Japan, Ltd.; Tokyo Electric Power Company; Terumo Company; and Miraka Holdings. The other authors declare no conflicts of interest.

Acknowledgments

We sincerely appreciate the members of the Suita Medical Foundation and the Suita City Health Center. We thank all researchers and medical staff in the Department of Preventive Cardiology, National Cerebral and Cardiovascular Center, for their excellent medical examinations and follow-up surveys. We also thank Satsuki-Junyukai, the society members of the Suita study. We appreciate the comments and suggestions from four anonymous reviewers that improved this manuscript significantly.

This study was supported by Health and Labour Sciences Research Grants (H27-Junkankitou-Ippan-009, H29-Junkankitou-Ippan-003, H30-Junkankitou-Ippan-005), the Intramural Research Fund of the National Cerebral and Cardiovascular Center (27-4-3), and a Grant-in-Aid for Scientific Research B (No. 16H05252).

References

1) World Health Organization. The atlas of heart disease and stroke. World Health Organization, Geneva, 2004
2) Vital Statistics of Japan, 2010, Tokyo. Health and Welfare Statistical Association 2010
3) D'Agostino RB Sr, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, Kannel WB. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation, 2008; 117: 743-753
4) Nishimura K, Okamura T, Watanabe M, Nakai M, Takegami M, Higashiyama A, Kokubo Y, Okayama A, Miyamoto Y. Predicting Coronary Heart Disease Using Risk Factor Categories for a Japanese Urban Population, and Comparison with the Framingham Risk Score: The Suita Study. J Atheroscler Thromb, 2014; 21: 784-798
5) Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, Brindle P. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. BMJ, 2008; 336: 1475-1482
6) Conroy RM, Pyörälä K, Fitzgerald AP, Sans S, Menotti A, De Backer G, De Bacquer D, Ducimetière P, Jousilahti P, Keil U, Njølstad I, Oganov RG, Thomsen T, Tunstall-Pedoe H, Tverdal A, Wedel H, Whincup P, Wilhelmsen L, Graham IM. SCORE project group. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. Eur Heart J, 2003; 24: 987-1003
7) Woodward M, Brindle P, Tunstall-Pedoe H. SIGN group

on risk estimation. Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). Heart, 2007; 93: 172-176
8) Siontis GC, Tzoulaki I, Siontis KC, Ioannidis JP. Comparisons of established risk prediction models for cardiovascular disease: systematic review. BMJ, 2012; 24: 344
9) Stout KK, Daniels CJ, Aboulhosn JA, Bozkurt B, Broberg CS, Colman JM, Crumb SR, Dearani JA, Fuller S, Gurvitz M, Khairy P, Landzberg MJ, Saidi A, Valente AM, Van Hare GF. 2018 AHA/ACC Guideline for the Management of Adults With Congenital Heart Disease: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol, 2018; pii: S0735-1097:36845-1
10) Ueshima H, Sekikawa A, Miura K, Turin TC, Takashima N, Kita Y, Watanabe M, Kadota A, Okuda N, Kadowaki T, Nakamura Y, Okamura T. Cardiovascular disease and risk factors in Asia: a selected review. Circulation, 2008; 118: 2702-2709
11) Ueshima H. Explanation for the Japanese Pradox: Prevention of increase in coronary heart disease and reduction in stroke. J Atheroscler Thromb, 2007; 14: 278-286
12) Arima H, Yonemoto K, Doi Y, Ninomiya T, Hata J, Tanizaki Y, Fukuhara M, Matsumura K, Iida M, Kiyohara Y. Development and validation of a cardiovascular risk prediction model for Japanese: the Hisayama study. Hypertens Res, 2009; 32: 1119-1122
13) Honda T, Yoshida D, Hata J, Hirakawa Y, Ishida Y, Shibata M, Sakata S, Kitazono T, Ninomiya T. Development and validation of modified risk prediction models for cardiovascular disease and its subtypes: The Hisayama Study. Atherosclerosis, 2018; 279: 38-44
14) NIPPON DATA80 Research Group. Risk assessment chart for death from cardiovascular disease based on a 19-year follow-up study of a Japanese representative population. Circ J, 2006; 70: 1249-1255
15) Shiba N, Nochioka K, Miura M, Kohno H, Shimokawa H. CHART-2 Investigators. Trend of westernization of etiology and clinical characteristics of heart failure patients in Japan--first report from the CHART-2 study. Circ J, 2011; 75: 823-833
16) Ministry of Internal Affairs and Communications, Statistics Bureau, Available at:https://www.stat.go.jp/data/kokusei/2010/final/pdf/01-02.pdf (Access Nov 6,2019)
17) Journal of Health and Welfare Statistics (In Japanese), Health, Labour and Welfare Statistics Association, 2013/2014,60,101
18) McClure LA, Kleindorfer DO, Kissela BM, Cushman M, Soliman EZ, Howard G. Assessing the performance of the Framingham Stroke Risk Score in the reasons for geographic and racial differences in stroke cohort. Stroke, 2014; 45: 1716-1720
19) Kokubo Y, Kamide K, Okamura T, Watanabe M, Higashiyama A, Kawanishi K, Okayama A, Kawano Y. Impact of high-normal blood pressure on the risk of cardiovascular disease in a Japanese urban cohort: the Suita study. Hypertension, 2008; 52: 652-659
20) Okamura T, Kokubo Y, Watanabe M, Higashiyama A, Miyamoto Y, Yoshimasa Y, Okayama A. Low-density lipoprotein cholesterol and non-high-density lipoprotein cholesterol and the incidence of cardiovascular disease in an urban Japanese cohort study: The Suita study. Atherosclerosis, 2009; 203: 587-592
21) Okamura T, Kokubo Y, Watanabe M, Higashiyama A, Ono Y, Miyamoto Y, Yoshimasa Y, Okayama A. Triglycerides and non-high-density lipoprotein cholesterol and the incidence of cardiovascular disease in an urban Japanese cohort: the Suita study. Atherosclerosis, 2010; 209: 290-294
22) Kokubo Y, Nakamura S, Okamura T, Yoshimasa Y, Makino H, Watanabe M, Higashiyama A, Kamide K, Kawanishi K, Okayama A, Kawano Y. Relationship between blood pressure category and incidence of stroke and myocardial infarction in an urban Japanese population with and without chronic kidney disease: the Suita Study. Stroke, 2009; 40: 2674-2679
23) Higashiyama A, Okamura T, Ono Y, Watanabe M, Kokubo Y, Okayama A. Risk of smoking and metabolic syndrome for incidence of cardiovascular disease--comparison of relative contribution in urban Japanese population: the Suita study. Circ J, 2009; 73: 2258-2263
24) Watanabe M, Kokubo Y, Higashiyama A, Ono Y, Okayama A, Okamura T. New diagnosis criteria for diabetes with hemoglobin A1c and risks of macro-vascular complications in an urban Japanese cohort: the Suita study. Diabetes Res Clin Pract, 2010; 88: e20-3
25) Kinoshita M, Yokote K, Arai H, Iida M, Ishigaki Y, Ishibashi S, Umemoto S, Egusa G, Ohmura H, Okamura T, Kihara S, Koba S, Saito I, Shoji T, Daida H, Tsukamoto K, Deguchi J, Dohi S, Dobashi K, Hamaguchi H, Hara M, Hiro T, Biro S, Fujioka Y, Maruyama C, Miyamoto Y, Murakami Y, Yokode M, Yoshida H, Rakugi H, Wakatsuki A, Yamashita S; Committee for Epidemiology and Clinical Management of Atherosclerosis. Japan Atherosclerosis Society (JAS) Guidelines for Prevention of Atherosclerotic Cardiovascular Diseases 2017. J Atheroscler Thromb, 2018; 25: 846-984
26) Shimamoto K, Ando K, Fujita T, Hasebe N, Higaki J, Horiuchi M, Imai Y, Imaizumi T, Ishimitsu T, Ito M, Ito S, Itoh H, Iwao H, Kai H, Kario K, Kashihara N, Kawano Y, Kim-Mitsuyama S, Kimura G, Kohara K, Komuro I, Kumagai H, Matsuura H, Miura K, Morishita R, Naruse M, Node K, Ohya Y, Rakugi H, Saito I, Saitoh S, Shimada K, Shimosawa T, Suzuki H, Tamura K, Tanahashi N, Tsuchihashi T, Uchiyama M, Ueda S, Umemura S; Japanese Society of Hypertension Committee for Guidelines for the Management of Hypertension. The Japanese Society of Hypertension Guidelines for the Management of Hypertension (JSH 2014). Hypertens Res, 2014; 37: 253-390
27) Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem, 1972; 18: 499-502
28) The World Health Organization MONICA Project (monitoring trends and determinants in cardiovascular disease): a major international collaboration. WHO MONICA Project Principal Investigators. J Clin Epidemiol, 1988; 41: 105-114
29) Walker AE, Robins M, Weinfeld FD. The national survey

of stroke. Clinical findings. Stroke, 1981; 12: I13-144
30) D'Agostino RB Sr, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, Kannel WB. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation, 2008; 117: 743-753
31) Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med, 1996; 15: 361-387
32) Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. Stat Med, 2004; 23: 2109-2123
33) Newson R. Comparing the predictive powers of survival models using Harrell's C or Somers' D. Stata Journal, 2010; 10: 339-358
34) Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr. Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med, 2008; 27: 157-172
35) Pencina MJ, D'Agostino RBS, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. StatMed, 2011; 30: 11-21
36) Goldman N, Glei DA, Weinstein M. What Matters Most for Predicting Survival? A Multinational Population-Based Cohort Study. PLoS One, 2016; 11
37) Saito I, Yamagishi K, Kokubo Y, Yatsuya H, Iso H, Sawada N, Inoue M, Tsugane S. Non-High-Density Lipoprotein Cholesterol and Risk of Stroke Subtypes and Coronary Heart Disease: The Japan Public Health Center-Based Prospective (JPHC) Study. J Atheroscler Thromb, 2020; 27: 363-374
38) Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke, 1991; 22: 983-988
39) Honda T, Yoshida D, Hata J, Hirakawa Y, Ishida Y, Shibata M, Sakata S, Kitazono T, Ninomiya T. Development and validation of modified risk prediction models for cardiovascular disease and its subtypes: The Hisayama Study. Atherosclerosis, 2018; 279: 38-44
40) Hata J, Ninomiya T, Hirakawa Y, Nagata M, Mukai N, Gotoh S, Fukuhara M, Ikeda F, Shikata K, Yoshida D, Yonemoto K, Kamouchi M, Kitazono T, Kiyohara Y. Secular trends in cardiovascular disease and its risk factors in Japanese: half-century data from the Hisayama Study (1961-2009). Circulation, 2013; 128: 1198-1205
41) Vliegenthart R, Oudkerk M, Hofman A, Oei HH, van Dijck W, van Rooij FJ, Witteman JC. Coronary calcification improves cardiovascular risk prediction in the elderly. Circulation, 2005; 112: 572-577

Supplemental Table 1. Univariable and multivariable hazard ratios among blood pressure categories by models

	Univariable adjusted for age	
	HR (95%CI)	*P*-value
SBP < 120 mmHg and DBP < 80 mmHg	0.407 (0.268, 0.617)	< 0.001
SBP = 120-139 mmHg or DBP = 80-89 mmHg	0.611 (0.414, 0.902)	0.013
SBP = 140-159 mmHg or DBP = 90-99 mmHg	0.816 (0.545, 1.222)	0.324
SBP 160 mmHg or DBP 90 mmHg	1.169 (0.760, 1.799)	0.477
(SBP < 140 or DBP < 90) and in medication	1.00	
(SBP ≥ 140 or DBP ≥ 190) and in medication	1.130 (0.754, 1.695)	0.554

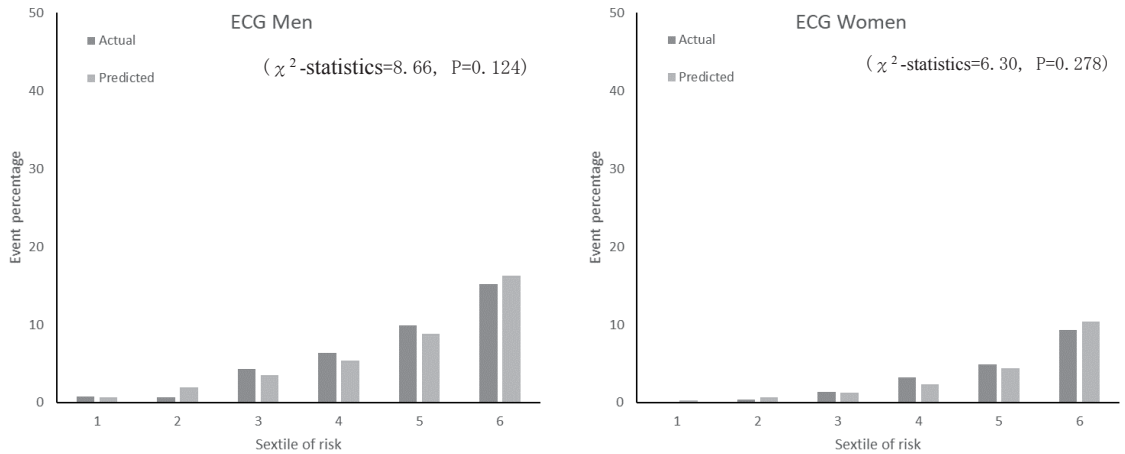
	Multivariable with Electrocardiogram		Multivariable without Electrocardiogram	
	HR (95%CI)	*P*-value	HR (95%CI)	*P*-value
SBP < 120 mmHg and DBP < 80 mmHg	0.412 (0.271, 0.626)	< 0.001	0.412 (0.271, 0.626)	< 0.001
SBP = 120-139 mmHg or DBP = 80-89 mmHg	0.601 (0.406, 0.890)	0.011	0.593 (0.401, 0.878)	0.009
SBP = 140-159 mmHg or DBP = 90-99 mmHg	0.803 (0.534, 1.208)	0.293	0.796 (0.530, 1.196)	0.272
SBP 160 mmHg or DBP 90 mmHg	1.141 (0.740, 1.759)	0.552	1.141 (0.740, 1.760)	0.551
(SBP < 140 or DBP < 90) and in medication	1.00		1.00	
(SBP ≥ 140 or DBP ≥ 190) and in medication	1.080 (0.718, 1.624)	0.712	1.093 (0.727, 1.623)	0.669

Multivariable model was adjusted for age, systolic blood pressure/diastolic blood pressure, Low density lipoprotein/ Non- high-density lipoprotein, high-density lipoprotein, smoking, diabetes mellitus and urinary protein. When each value of blood pressures (SBP and DBP) or lipid profiles (LDL-C and non-HDL-C) belonged to different categories, we adopted the higher category to estimate hazard ratios.

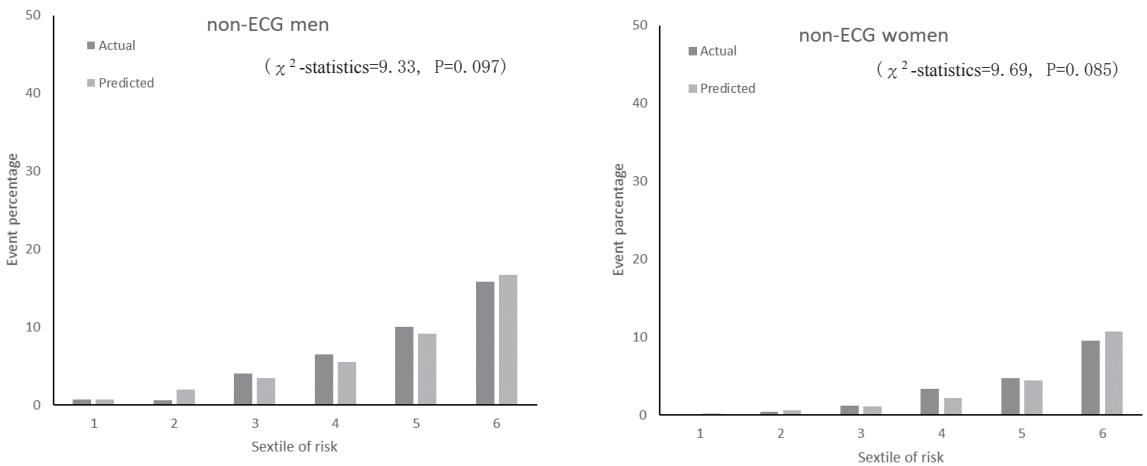
Supplemental Table 2. Number of incident cases and incident rates by each categorical variable

	Number of participants	Number of incident cases	Persons years	Incident rate/1,000 persons years
Men	3,080	351	44,338	7.92
Women	3,470	241	55,335	4.36
Age, year				
30-39	1,087	15	17,167	0.87
40-49	1,303	52	22,875	3.03
50-59	1,551	122	25,449	5.33
60-64	920	108	13,968	7.73
65-69	621	103	8,704	11.83
70-74	682	121	7,923	15.27
75-79	386	71	3,588	19.79
SBP < 120 mmHg and DBP < 80 mmHg	2,456	100	40,025	2.50
SBP = 120-139 mmHg or DBP = 80-89 mmHg	2,113	177	32,706	5.41
SBP = 140-159 mmHg or DBP = 90-99 mmHg	881	112	12,556	8.92
SBP ≥ 160 mmHg or DBP ≥ 100 mmHg or in-medication	1,100	203	14,387	14.11
Non-HDL-C < 170 mg/dL and LDL-C < 140 mg/dL	4,019	311	60,714	5.12
Non-HDL -C ≥ 170 mg/dL or LDL-C ≥ 140 mg/dL	2,531	281	38,959	7.21
HDL-C < 40 mg/dL	944	129	14,009	9.21
HDL-C = 40-59 mg/dL	3,442	326	52,403	6.22
HDL-C ≥ 60 mg/dL	2,164	137	33,262	4.12
Smoking	1,962	199	28,743	6.92
Non-Smoking	4,588	393	70,931	5.54
Diabetes mellitus	326	67	4,064	16.49
Non- Diabetes mellitus	6,224	525	95,609	5.49
Urinary protein 1 + or more	1,208	150	17,355	8.64
Urinary protein less than 1 +	5,342	442	82,318	5.37
Atrial fibrillation	40	10	401	24.94
No Atrial fibrillation	6,510	582	99,272	5.86
Left ventricular hypertrophy	120	28	1,456	19.23
No Left ventricular hypertrophy	6,430	564	98,217	5.74
Stroke	6,550	341	99,673	3.42
Cerebral Infarction	6,550	206	99,673	2.07
Intracerebral hemorrhage	6,550	67	99,673	0.67
Subarachnoid Hemorrhage	6,550	30	99,673	0.30
Definite MI with MONICA criteria	6,550	75	99,673	0.75
Possible MI with MONICA criteria	6,550	22	99,673	0.22
Coronary Intervention	6,550	93	99,673	0.93
Sudden cardiac death	6,550	4	99,673	0.04
MI death with death certificate only	6,550	57	99,673	0.57
All CHDs (MI with MONICA criteria and other CHDs)	6,550	251	99,673	2.52

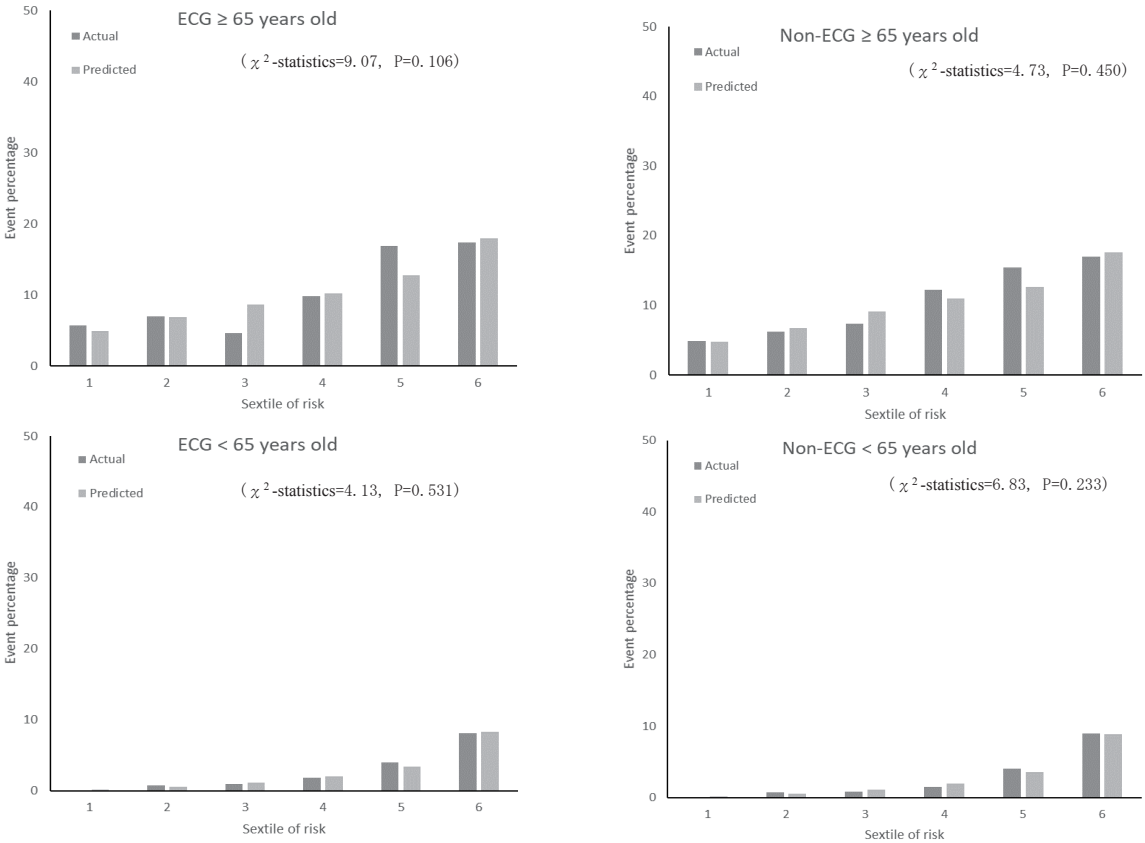
MI: myocardial infarction CHD: Coronary Heart Disease



Supplemental Fig. 1a. Actual and predicted cardiovascular event percentage by sex and sextiles of risk with ECG



Supplemental Fig. 1b. Actual and predicted cardiovascular event percentage by sex and sextiles of risk with non-ECG

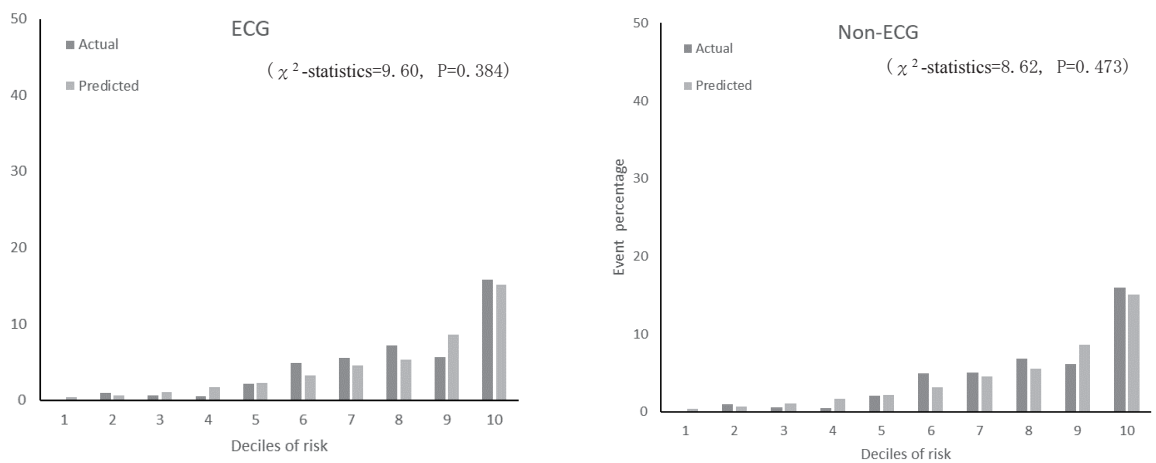


Supplemental Fig. 1c. Actual and predicted cardiovascular event percentage by age category (<65, ≥ 65) and sextiles of risk with ECG and non-ECG

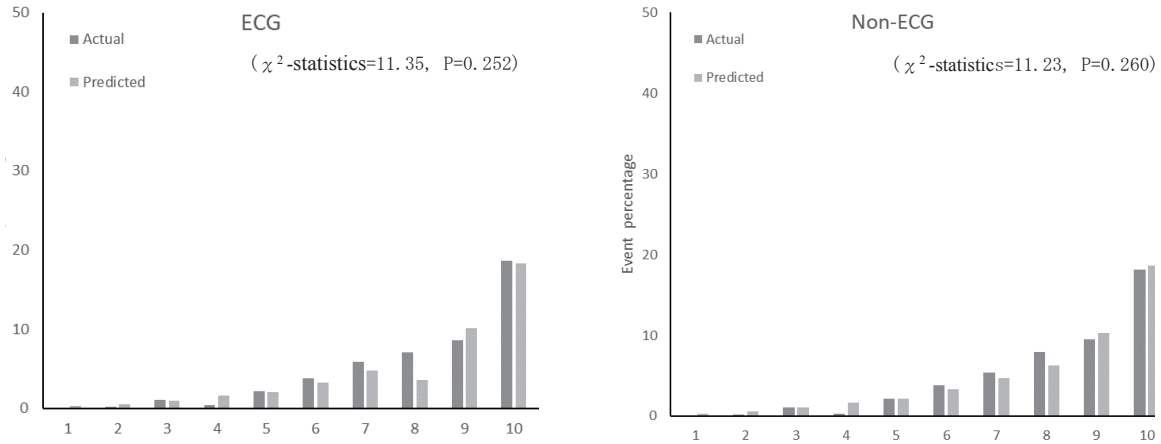
Supplement Table 3. Cardiovascular disease risk points and regression coefficient with/ without Electrocardiogram using derivation dataset

		Model with Electrocardiogram		Model without Electrocardiogram	
		Coefficient	Score	Coefficient	Score
Sex (Men)		0.26	3	0.26	3
Age	30-39 years	0	0	0	0
	40-49 years	0.81	8	0.82	8
	50-59 years	1.36	14	1.38	14
	60-64 years	1.71	17	1.74	18
	65-69 years	2.17	22	2.20	22
	70-74 years	2.50	25	2.54	25
	75-79 years	2.66	27	2.71	27
BP	SBP < 120 mmHg and DBP < 80 mmHg	−0.51	−5	−0.48	−5
	SBP = 120-139 mmHg or DBP = 80-89 mmHg	0	0	0	0
	SBP = 140-159 mmHg or DBP = 90-99 mmHg	0.24	2	0.24	2
	SBP ≥ 160 mmHg or DBP ≥ 100 mmHg or in medication	0.57	6	0.58	6
Non-HDL-C < 170 mg/dL and LDL-C < 140 mg/dL		0	0	0	0
Non-HDL-C ≥ 170 mg/dL or LDL-C ≥ 140 mg/dL		0.18	2	0.17	2
HDL-C	HDL-C < 40 mg/dL	0	0	0	0
	HDL-C = 40-59 mg/dL	−0.30	−3	−0.30	−3
	HDL-C ≥ 60 mg/dL	−0.47	−5	−0.46	−5
Smoking		0.38	4	0.36	4
DM		0.62	6	0.64	6
Urinary protein 1 + or more		0.24	2	0.26	3
Atrial fibrillation		1.12	11		
Left ventricular hypertrophy		0.54	5		

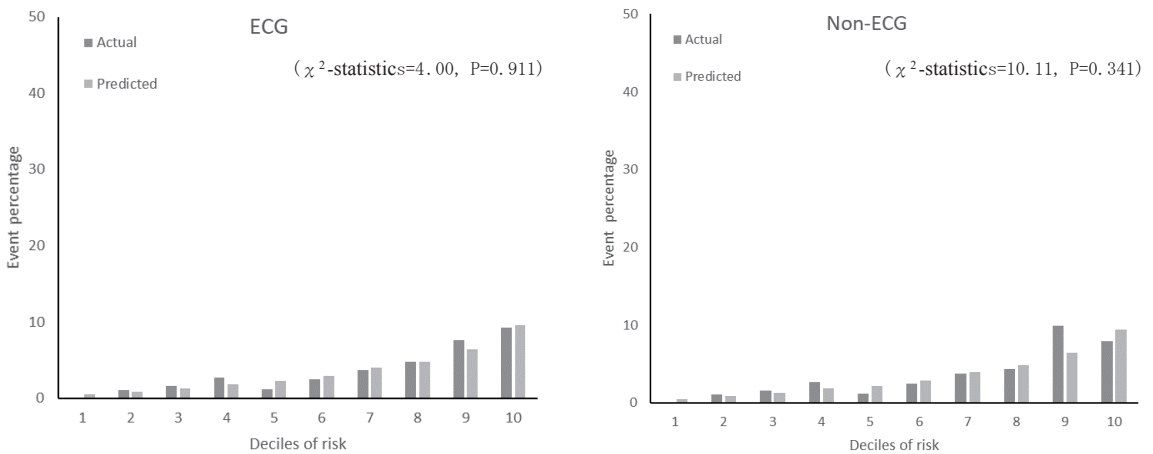
When each value of blood pressures (SBP and DBP) or lipid profiles (LDL-C and non-HDL-C) belonged to different categories, we adopted the higher category to estimate regression coefficients and scores.



Supplemental Fig. 2. Actual and predicted cardiovascular event percentage by deciles of risk with ECG and non-ECG using validation dataset



Supplemental Fig. 3a. Actual and predicted cardiovascular event percentage by deciles of risk with ECG and non-ECG using first cohort dataset



Supplemental Fig. 3b. Actual and predicted cardiovascular event percentage by deciles of risk with ECG and non-ECG using second cohort dataset