Validity of a Risk Prediction Equation for CKD After 10 -Years of Follow-up in a Japanese Population: The Ibaraki Prefectural Health Study

Mitsumasa Umesawa, MD, PhD[1,2], Toshimi Sairenchi, PhD[1,2], Yasuo Haruyama, MD, PhD[1], Masanori Nagao, PhD[1,2], Kazumasa Yamagishi, MD, PhD[2,3], Fujiko Irie, MD, PhD[2,4], Hiroshi Watanabe, MD[5], Gen Kobashi, MD, PhD[1,2], Hiroyasu Iso, MD, PhD[6], Hitoshi Ota, MD, PhD[2]

[1] Department of Public Health, Dokkyo Medical University, Mibu, Japan

[2] Ibaraki Health Plaza, Ibaraki Health Service Association, Mito, Japan

[3] Department of Public Health Medicine, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan

[4] Department of Health and Welfare, Ibaraki Prefectural Office, Mito, Japan

[5] Ibaraki Health Service Association, Mito, Japan

[6] Public Health, Department of Social Medicine, Osaka University Graduate School of Medicine, Suita, Japan

Corresponding author

Mitsumasa Umesawa

Department of Public Health, Dokkyo Medical University, Mibu 321-0293, Japan

Tel.: +81-282-87-2133

Fax: +81-282-86-2935

Email: umesawa@dokkyomed.ac.jp

Word count (main text): 2846

Word count (abstract): 319

Number of tables: 4

Number of figures: 2

Short title: CKD risk prediction in a Japanese population

Abstract

Background: Chronic kidney disease (CKD) is an important health problem for which risk equations have been developed for Western populations. This study aimed to develop and validate a risk prediction equation for CKD in a Japanese population.

Study Design: Observational cohort study.

Setting & Participants: The study included 135 007 participants who completed an annual health check-up between 1993 and 1996 in the Ibaraki Prefecture in Japan. Participants were initially free of CKD stage 3, 4 or 5, and proteinuria 2+ or 3+. Follow-up information was available from health check-ups 10 years after initial evaluation. We used data from 17 892 men and 40 963 women in the northern region of the prefecture for the development of risk prediction equations, and 22 992 men and 49 952 women in the southern region for external validation.

Predictors: Age, estimated glomerular filtration rate (eGFR), body mass index, proteinuria, hematuria, hypertension, diabetes mellitus, smoking, and drinking.

Outcome: Occurrence of eGFR <60 mL/min/1.73 m^2 and/or proteinuria 2+ or 3+.

Analytical Approach: Logistic regression analysis to estimate risk for CKD stratified by sex.

Results: During follow-up, 7500 cases of CKD developed in the northern region and 8964 in the southern region. Older age, proteinuria(+), higher systolic blood pressure, medication for hypertension, and current smoking were associated with the risk of CKD in both sexes while higher eGFR and daily alcohol intake were associated with lower risk. C-statistics of the risk estimation equations for CKD at 10 years were >0.8 for both the development and external validation populations, and discrimination of the risk estimation was fairly good in men and women.

Limitations: Fluctuations in variables were not evaluated because the study used annual health check-ups. This study excluded a large number of people for whom a 10-year health check-up was not available.

Conclusions: Estimations of risk for CKD after 10 years of follow-up in a general Japanese population can be achieved with a high level of validity.

Index words: estimated glomerular filtration rate, chronic kidney disease, risk score, Japanese, Asian, cohort study

Summary of article: Chronic kidney disease (CKD) is an important health problem. A risk estimation tool (risk score for CKD prevention) has been developed for Western populations, but not Asian populations. This study aimed to develop and validate prediction equations for CKD at 10 years in a Japanese population. We included 135 007 Japanese people who had completed annual health check-ups between 1993 and 1996. Participants were free from CKD (defined as CKD stage 3 or higher and/or proteinuria 2+ or 3+) and/or renal disease, and information was available from their health check-ups 10 years later. We included 58 855 people living in the northern region for development of prediction equations. The remaining participants were included in the external validation. The discrimination of the risk prediction equations was fairly good in men and women. The c-statistics for external validation were >0.8.

Introduction

Chronic kidney disease (CKD) is one of the most common non-communicable diseases. CKD is associated with the risk for end-stage renal disease (ESRD), mortality from cardiovascular disease (CVD), and mortality from any causes (1). Prevention of CKD is important, but it is difficult to identify patients with early stage CKD because it is asymptomatic.

A risk prediction tool such as risk score is useful to evaluate whether people are at high risk for developing chronic diseases. For example, there are several risk scores for coronary heart disease and CVD (2, 3). For CKD, a major risk score for Western populations was developed from the results of the Framingham Heart Study and Atherosclerosis Risk in Communities (4). However, in Asian populations, there are only a few prospective studies of Japanese people that have examined the association between risk factors and the risk for CKD (5, 6), and a risk prediction tool for CKD for Asian general populations has not been developed.

According to the Japanese Society of Nephrology, the prevalence of CKD in Japanese adults is approximately 20% (7). This number is expected to increase because of the aging Japanese population. Therefore, it is important to develop a useful risk prediction tool to detect people at high risk for CKD.

The aim of this study was to develop and validate risk prediction equations for CKD using those factors generally measured at annual health check-ups in Japan. A predictive tool based on data collected from the general population may be useful in both public health and medical settings.

Methods

Study participants

Participants were Japanese people aged 40–74 years old who were living in Ibaraki Prefecture. A baseline survey was conducted between 1993 and 1996 during which time a total of 307 531 people (105 381 men and 202 150 women) attended annual health check-ups. Of these, 141 608 also attended an annual health check-up 10 years later. We included 135 007 people (41 002 men and 94 005 women) because they had completed the examination and were free from renal disease or CKD at the first check-up. We defined CKD as estimated glomerular filtration rate (eGFR) <60 mL/min/1.73 m^2 and/or proteinuria 2+ or 3+.

Of the 135 007 participants included, we used data from 58 855 people (17 892 men and 40 963 women) living in the northern region of Ibaraki Prefecture for the development of the risk prediction equations. For validation of the risk prediction equations, we used data from 76 152 people (23 110 men and 53 042 women) living in the southern region of Ibaraki Prefecture for external validation.

Informed consent was obtained from community representatives to conduct an epidemiological study on the basis of the guidelines of the Council for International Organizations of Medical Science. The study was approved by the Ethics Committee of Ibaraki Prefecture.

Risk factor survey

The annual health check-up was performed by the Ibaraki Health Service Association. Health check-ups conducted between 1993 and 1996 were regarded as the baseline survey. Health check-ups conducted 10 years after the baseline survey were regarded as the follow-up survey. If a person had undergone an annual health check-up twice or more between 1993 and 1996, the earliest data was used for the analyses. The check-up consisted of body weight measurement with light clothing, medical history, alcohol/smoking status, blood pressure measurement, blood tests, and dipstick urine tests. Serum creatinine was measured using the

modified Jaffe method until March 2004 and using the enzymatic method from April 2004 with an autoanalyzer (Hitachi 7350; Hitachi, Tokyo, Japan or RX-20; JEOL, Tokyo, Japan). The following formula was used to calculate the creatinine value based on the Jaffe method value and the enzymatic method value: serum creatinine (enzymatic method) = 1.0085 × creatinine (Jaffe method) − 0.265 mg/dL. The creatinine values reported in this study represent the calculated enzymatic values. eGFR was calculated using the following formula established by the working group of the Japanese Chronic Kidney Disease Initiative: eGFR (mL/min/1.73 m^2) = 1.94 × (serum creatinine)$^{-1.094}$ × (age)$^{-0.287}$ × 0.739 (if female) (8). Blood glucose was measured using the autoanalyzer. Proteinuria and hematuria were tested with dipsticks (Ames Hemacombisticks; Bayer-Sankyo, Tokyo, Japan). Blood pressure was measured using standard mercury sphygmomanometers or automated sphygmomanometers on the right arm. If blood pressure was high with the first measurement, the blood pressure measurement was repeated once.

Statistical analysis

Multivariable logistic regression models were used to evaluate risk factors for CKD. Regression coefficients, odds ratios (OR), and their 95% confidence intervals were calculated after adjustment for age and risk factors. The risk factors were eGFR (mL/min/1.73 m^2), dipstick proteinuria (+ and missing), dipstick hematuria (+, 2+, 3+, and missing), body mass index (kg/m^2), systolic blood pressure (mmHg), medication for hypertension, medication for diabetes mellitus, smoking status (past smoker and current smoker), and alcohol intake (occasional and every day). We also included glucose tolerance as a risk factor. Fasting status was self-reported and the oral glucose tolerance testing was not performed. We used fasting serum glucose (mg/dL) or non-fasting serum glucose (mg/dL) for the index of glucose tolerance. For proteinuria, hematuria, medication for hypertension, medication for diabetes

mellitus, and smoking and alcohol intake, we used participants who did not have the risk factor as the reference. For systolic blood pressure and glucose tolerance, if measurements were not taken for a participant, it was regarded as missing.

Two models were developed for the risk prediction equations. The simple risk prediction equation included age, eGFR, proteinuria, and hematuria. The full risk prediction equation included age, eGFR, proteinuria, hematuria, body mass index, systolic blood pressure, medication for hypertension, glucose tolerance, medication for diabetes mellitus, and smoking and alcohol intake. The risk predicted equations for the logistic model were based on the formula: $\hat{p} = 1 / (1 + \exp(-(\text{intercept} + \text{sum of the estimate per risk factors}))$. The estimate of each risk factor was calculated as: (regression coefficient) × (value of each factor). We used risk factors that were significantly associated with the risk for CKD at 10 years for calculation of the risk estimate. We developed a calibration plot for each model to examine the discrimination among risk scores (Figures 1 and 2). We also calculated the c-statistic for each model to validate the risk prediction equations for CKD at 10 years. Risk scores were calculated using the Framingham Study risk score functions (9). The following classifications were used. eGFR: 60–74, 75–89, 90–119, or ≥120 mL/min/1.73 m^2. Blood pressure: stage 1 hypertension: 140 mmHg ≤ systolic blood pressure < 160 mmHg and/or 90 mmHg ≤ diastolic blood pressure < 100 mmHg; stage 2 hypertension: 160 mmHg ≤ systolic blood pressure < 180 mmHg and/or 100 mmHg ≤ diastolic blood pressure < 110 mmHg; stage 3 hypertension: 180 mmHg ≤ systolic blood pressure and/or 110 mmHg ≤ diastolic blood pressure, medicated and missing. Glucose tolerance: impaired glucose tolerance: 110 mg/dL ≤ fasting serum glucose < 126 mg/dL or 140 mg/dL ≤ non-fasting serum glucose < 200 mg/dL. Diabetes: 126 mg/dL ≤ fasting serum glucose or 200 mg/dL ≤ non-fasting serum glucose, medicated and missing. These classifications were used instead of continuous variables to simplify the calculations. We excluded body mass index from risk score because

body mass index was not significantly associated with the risk of CKD at 10 years in this analysis. The methods for how the risk scores were determined are provided in Supplemental Table 1. We determined 1 point as the risk associated with a 5-year increase in age. We defined the constant as: the estimate per 1-year increment of the regression coefficients of the multiple regression model multiplied by 5. We determined points associated with each category by dividing the estimate of each risk factor by the constant. The points of each risk factor were rounded to the nearest integer. Estimated risk for CKD associated with point totals was based on the Framingham Study risk score functions (9). For Model 1 (simple risk score), risk score comprised age, proteinuria, hematuria, and baseline eGFR. For Model 2 (full risk score), risk score comprised the variables in Model 1 plus blood pressure, glucose tolerance, smoking status, and drinking status. Estimated risk for CKD was calculated as: $1 / (1 + \exp(4.1537 - 0.268 \times \text{points}))$ for men and $1 / (1 + \exp(3.7942 - 0.121 \times \text{points}))$ for women in Model 1; and $1 / (1 + \exp(4.2248 - 0.257 \times \text{points}))$ for men and $1 / (1 + \exp(3.8008 - 0.1055 \times \text{points}))$ for women in Model 2. SAS software (ver. 9.4; SAS Institute Inc., Cary, NC, USA) was used for all analyses. A *P*-value <0.05 was considered statistically significant (two-tailed analysis).

Results

Table 1 shows the characteristics of participants stratified by geographic area. During the 10-year follow-up, a total of 16 464 people developed CKD. In the northern region, 2840 men and 4660 women developed CKD, and in the southern region, 3266 men and 5698 women developed CKD.

Table 2 and 3 shows the regression coefficients and the multivariable-adjusted OR for CKD at 10 years among people living in the northern region.

For Model 1, age at baseline and proteinuria (+) were positively, and eGFR was negatively, associated with the risk for CKD in men and women. Weak (+) hematuria was positively associated with the risk for CKD in men but not in women, while severe (3+) hematuria was positively associated with the risk for CKD in women but not in men.

For Model 2, associations between risk factors included in Model 1 and the risk for CKD was consistent. Systolic blood pressure, medication for hypertension, and current smoker were positively associated with the risk for CKD in both men and women. Fasting serum glucose, non-fasting serum glucose, and daily alcohol intake were negatively associated with the risk for CKD in both men and women. Body mass index was positively, and past smoking was negatively associated with the risk for CKD in men but not women. Medication for diabetes mellitus was positively, and occasional alcohol intake was negatively, associated with the risk for CKD in women but not men. According to these results, we developed risk predictive equations both simple and full model (Table 4).

Figure 1 shows the calibration plot for each model in men. The discrimination of the prediction equations were fairly good until 70% of predicted probabilities. Figure 2 shows the calibration plot for each model in women. The discrimination of prediction equations were fairly good until 60% of predicted probabilities.

We also calculated the c-statistics for each risk prediction equation. For Model 1, the c-statistic of the 10-year risk for CKD in people living in the southern region (for external validity) was 0.827 for men and 0.814 for women. For people living in the northern region (for develpoment), the c-statistic was 0.823 for men and 0.824 for women. For Model 2, the c-statistic was 0.831 and 0.815 for men and women living in the southern region and 0.827 and 0.826 for men and women living in the northern region, respectively.

Based on these results, we developed the risk score for each category of risk factors for CKD at 10 years (Table S2). The estimated risk for CKD at 10 years is shown in Supplemental Table 3, according to the developed risk score.

Discussion

We developed and validated risk prediction equations for CKD at 10 years based on annual health check-ups among a Japanese general population. The risk prediction equations showed good validity for both the development and external validation. According to the risk estimates, baseline eGFR value was the most important predictor of CKD at 10 years, and dipstick proteinuria was the second most important predictor.

In the current study, the association between diabetes mellitus and the risk for CKD was different from some prior studies (4, 10-12). Fasting serum glucose and non-fasting serum glucose were negatively associated with the risk for CKD at 10 years, while medication for diabetes mellitus was positively associated with the risk for CKD at 10 years in women. In the additional analysis (Table S4), impaired glucose tolerance was negatively associated, and diabetes mellitus without medication was not associated, with the risk for CKD at 10 years. We suggest that the differences between the previous studies and the current study are related to the progressive nature of CKD in patients with diabetes (13). In general, the eGFR of patients with diabetes often remains stable or even reaches hyperfiltration in 10–15 years. Therefore, we suspect that the lack of association between non-medicated diabetes mellitus or impaired glucose tolerance with elevated risk for CKD in this cohort of Japanese participants is because diabetic renal injury first manifests as hyperfiltration before loss of eGFR becomes evident.

In the current study, we included dipstick hematuria and smoking and alcohol intake as potential risk factors, although they were not used in the risk score for the Western

population (4), because previous epidemiological studies have reported an association between those factors and the risk for CKD or ESRD (5, 6, 12, 14). For hematuria, a prospective study of Japanese people reported a positive association between hematuria and the risk for CKD (5). In the study, weak (+) hematuria and moderate to severe (2+ and 3+) hematuria were positively associated with the risk for CKD stage 3 or higher. A prospective study of Israeli men and women reported that microscopic hematuria was positively associated with the risk for ESRD (14). In the current study, the association between dipstick hematuria and the risk for CKD was inconsistent between men and women. Regarding smoking, previous epidemiological studies report a positive association between smoking and the risk for CKD (12, 15), which is consistent with our results. A prospective study of Australian men and women reported an inverse association between alcohol intake and the risk for CKD (16), while another prospective study of American men and women reported a positive association between frequent or heavy alcohol intake and the risk for CKD (15). In the current study, daily alcohol intake was inversely associated with the risk for CKD in both men and women.

The strengths of the current study include the study design and the participants. We used a prospective cohort design with a large number of Japanese participants from the general population. This allowed us to develop and validate risk prediction equations that included many risk factors, which should help physicians evaluate the estimated risk for CKD in their patients. In addition, we developed CKD risk prediction equations using variables whose measurements are not expensive or difficult, meaning the prediction equations and risk score can be widely used. Evaluation of CKD risk in a health survey is beneficial for the detection of high risk CKD populations, and would allow for early intervention to prevent CKD progression.

This study had some limitations. First, the study was based on annual health check-ups. Examinations performed once a year do not allow for the evaluation of any fluctuation in serum creatinine and other variables. However, any errors concerning the misclassification were non-differential. Second, a large number of people were excluded from the analysis. The main reason for exclusion was that they did not attend an annual health check-up 10 years after the first. We did not have information for why they did not attend the health check-up; however, compared with people who attended the 10-year annual health check-up, people who did not attend were almost 3.6 years older for men and 4.4 years older for women at the baseline survey. Also, it seems likely that people who did not attend annual health check-up 10-years later might include those with severe illness, including CKD, and had begun regular medical examinations by a medical doctor between the baseline and the follow-up surveys. Our results may be affected by this sampling bias. Third, blood samples were not all obtained from participants under fasting conditions. We asked participants whether they had fasted or not at the interview, but misclassification may exist. However, we did examine the effect of diabetes on the risk for CKD and found that the effect was limited. Fourth, we defined CKD as eGFR <60 mL/min/1.73 m^2 and/or proteinuria (2+ or 3+). According to KDIGO 2012, CKD includes kidney damage as albuminuria, urine sediment abnormalities, electrolyte and other abnormalities due to tubular disorders, abnormalities detected by histology, structural abnormalities detected by imaging, and history of kidney transplantation. We could not include these classifications because we did not have information about them. Fifth, we could not use information about cystatin-C because its measurement was not common between 1993 and 1996 in Japan. Therefore, the accuracy of eGFR in the fine categories is not guaranteed. Sixth, we developed calibration plots for each model. According to the plot, the reliability of the risk prediction equations was limited among people whose estimated risk was high.

In conclusion, we developed risk prediction equations for CKD at 10 years of follow-up among a Japanese general population. The validity of the risk prediction equations was high. These risk prediction equations may help physicians to identify individuals at high risk for CKD, particularly in Asian ethnic groups.

Acknowledgments

The authors thank the staff and participants of the Ibaraki Prefectural Health Study.

Contributions: Research idea and study design: MU, TS, HI; data acquisition: MU, TS, HW, HI, HO; data analysis/interpretation: MU, TS, YH, MN, KY, FI, GK, HI; statistical analysis: MU, TS; supervision or mentorship: TS, YH, MN, KY, FI, HW, GK, HI, HO. Each author contributed important intellectual content during manuscript drafting or revision and accepts accountability for the overall work by ensuring that questions pertaining to the accuracy or integrity of any portion of the work are appropriately investigated and resolved.

Support and financial disclosure declaration

Support: None. There was no support/funding for this study.

Financial Disclosure: The authors declare no conflict of interest.

Peer Review: Received December 26, 2016. Evaluated by two external peer reviewers, a Statistics Editor, an Associate Editor, and the Editor-in-Chief. Accepted in revised form September 7, 2017.

References

1. Nitsch D, Grams M, Sang Y, et al. Associations of estimated glomerular filtration rate and albuminuria with mortality and renal failure by sex: a meta-analysis. *BMJ.* 2013;346:f324.
2. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation.* 1998;97(18):1837-1847.
3. 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(10):1249-1255.
4. O'Seaghdha CM, Lyass A, Massaro JM, et al. A risk score for chronic kidney disease in the general population. *Am J Med.* 2012;125(3):270-277.
5. Yamagata K, Ishida K, Sairenchi T, et al. Risk factors for chronic kidney disease in a community-based population: a 10-year follow-up study. *Kidney Int*. 2007;71(2):159-166.
6. Nishikawa K, Takahashi K, Okutani T, et al. Risk of chronic kidney disease in non-obese individuals with clustering of metabolic factors: a longitudinal study. *Intern Med*. 2015;54(4):375-382.
7. Imai E, Horio M, Iseki K, et al. Prevalence of chronic kidney disease (CKD) in the Japanese general population predicted by the MDRD equation modified by a Japanese coefficient. *Clin Exp Nephrol.* 2007;11(2):156-163.
8. Matsuo S, Imai E, Horio M, et al. Revised equations for estimated GFR from serum creatinine in Japan. *Am J Kidney Dis.* 2009;53(6):982-992.

9. Sullivan LM, Massaro JM, D'Agostino RB Sr. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Stat Med.* 2004;23(10):1631-1660.
10. Derakhshan A, Bagherzadeh-Khiabani F, Arshi B, Ramezankhani A, Azizi F, Hadaegh F. Different Combinations of Glucose Tolerance and Blood Pressure Status and Incident Diabetes, Hypertension, and Chronic Kidney Disease. *J Am Heart Assoc.* 2016;5(8):e003917.
11. Chang PY, Chien LN, Lin YF, Wu MS, Chiu WT, Chiou HY. Risk factors of gender for renal progression in patients with early chronic kidney disease. *Medicine (Baltimore).* 2016;95(30):e4203.
12. Fox CS, Larson MG, Leip EP, Culleton B, Wilson PW, Levy D. Predictors of new-onset kidney disease in a community-based population. *JAMA.* 2004;291(7):844-850.
13. Remuzzi G, Benigni A, Remuzzi A. Mechanisms of progression and regression of renal lesions of chronic nephropathies and diabetes. *J Clin Invest*. 2006;116(2):288-296.
14. Vivante A, Afek A, Frenkel-Nir Y, et al. Persistent asymptomatic isolated microscopic hematuria in Israeli adolescents and young adults and risk for end-stage renal disease. *JAMA.* 2011;306(7):729-736.
15. Shankar A, Klein R, Klein BE. The association among smoking, heavy drinking, and chronic kidney disease. *Am J Epidemiol.* 2006;164(3):263-271.
16. White SL, Polkinghorne KR, Cass A, Shaw JE, Atkins RC, Chadban SJ. Alcohol consumption and 5-year onset of chronic kidney disease: the AusDiab study. *Nephrol Dial Transplant*. 2009;24(8):2464-2472.

Table 1 Characteristics of Subjects Stratified by Sex and Area of Residence (Southern vs. Northern part of Ibaraki Prefecture, Japan) at Baseline

	Men			Women	
	Northern region	Southern region		Northern region	Southern region
N	17892	23110		40963	53042
Cases of chronic kidney disease at 10-years	2840	3266		4660	5698
Age (years)	58.5	56.8		55.0	54.3
Body mass index (kg/m^2)	23.4	23.5		23.4	23.4
eGFR (ml/min/1.73m^2)	87.9	89.3		95.1	96.9
Dipstick Proteinuria (%)					
+	1.4	1.2		0.5	0.8
Missing	0.6	0.5		0.9	1.0
Dipstick Hematuria (%)					
+	3.5	6.3		7.1	10.9
2+	1.4	2.9		4.1	6.8
3+	0.9	0.9		2.4	2.6
Missing	0.6	0.5		4.4	5.8
Blood pressure levels					
Systolic blood pressure (mmHg)	134.3	134.2		129.2	129.4
Diastolic blood pressure (mmHg)	80.3	81.1		76.9	77.4
Medicated for hypertension (%)	16.9	16.6		14.2	14.2
Glucose tolerance					
Fasting serum glucose (mg/dl)[a]	102.4	101.1		97.7	96.4
Non-fasting serum glucose (mg/dl)[b]	118.4	118.3		108.0	107.9
Medicated for diabetes mellitus	2.7	3.0		1.4	1.4
Smoking (%)					
Past smoker	30.8	27.0		0.4	0.6
Current smoker	44.4	49.1		3.9	5.0
Alcohol intake (%)					
Occasional	14.0	13.1		6.4	7.3
Everyday	53.9	58.1		3.8	4.8

eGFR: estimated glomerular filtration rate at baseline

[a] N=5952 in men living in northern region, N=9312 in men living in southern region, N=12715 in women living in northern region and N=21840 in women living in southern region

[b] N=11940 in men living in northern region, N=13798 in men living in southern region, N=28248 in women living in northern region and N=31202 in women living in southern region

Table 2　Regression coefficients and odds ratios of Chronic Kidney Disease (stage 3 or higher and/or having proteinuria 2+ or 3+) among men Living in Northern Ibaraki Prefecture, Japan				
Model 1 (simple model)				
Risk factor	Regression coefficient	odds ratios	95 percent confidence intervals	P value
Intercept	3.3627	-	-	<0.001
Age (years)	0.0404	1.041	1.035 - 1.048	<0.001
eGFR (ml/min/1.73m^2)	-0.0932	0.911	0.907 - 0.915	<0.001
Dipstick Proteinuria (%)				
+	1.5395	4.662	3.412 - 6.369	<0.001
Missing	-0.1553	0.856	0.468 - 1.566	0.6
Dipstick Hematuria (%)				
+	0.4237	1.528	1.225 - 1.905	<0.001
2+	0.2164	1.242	0.873 - 1.767	0.2
3+	0.3232	1.382	0.892 - 2.141	0.1
Missing	-	-	-	-
Model 2 (full model)				
Risk factor	Regression coefficient	odds ratios	95 percent confidence intervals	P value
Intercept	1.7047	-	-	<0.001
Age (years)	0.0402	1.041	1.034 - 1.048	<0.001
eGFR (ml/min/1.73m^2)	-0.0928	0.911	0.907 - 0.915	<0.001
Dipstick Proteinuria (%)				
+	1.4492	4.260	3.183 - 5.990	<0.001
Missing	-0.1805	0.835	0.451 - 1.516	0.6
Dipstick Hematuria (%)				
+	0.4124	1.510	1.209 - 1.883	<0.001
2+	0.1842	1.202	0.844 - 1.711	0.3
3+	0.2892	1.335	0.854 - 2.057	0.2
Missing	-	-	-	-
Body Mass Index (kg/m^2)	0.0323	1.033	1.016 - 1.050	<0.001
Systolic blood pressure (mmHg)	0.0077	1.008	1.005 - 1.011	<0.001
Medicated for hypertension	0.2428	1.275	1.139 - 1.427	<0.001
Glucose tolerance				
Fasting serum glucose (mg/dl)	-0.0017	0.998	0.997 - 1.000	0.05
Non-fasting serum glucose (mg/dl)	-0.0016	0.998	0.997 - 1.000	0.02

Medicated for diabetes mellitus	0.2078	1.231	0.945 - 1.604	0.1
Smoking status				
Past smoker	-0.1581	0.854	0.756 - 0.964	0.01
Current smoker	0.2313	1.260	1.123 - 1.415	<0.001
Alcohol intake				
Occasional	-0.0120	0.988	0.858 - 1.138	0.9
Everyday	-0.1639	0.849	0.767 - 0.939	0.002

eGFR: estimated glomerular filtration rate at baseline

Table 3 Regression coefficients and odds ratios of Chronic Kidney Disease (stage 3 or higher and/or having proteinuria 2+ or 3+) among women Living in Northern Ibaraki Prefecture, Japan				
Model 1 (simple model)				
Risk factor	Regression coefficient	odds ratios	95 percent confidence intervals	P value
Intercept	3.0796	-	-	<0.001
Age (years)	0.0309	1.031	1.027 - 1.036	<0.001
eGFR (ml/min/1.73m^2)	-0.0816	0.922	0.919 - 0.924	<0.001
Dipstick Proteinuria (%)				
+	1.1677	3.214	2.289 - 4.515	<0.001
Missing	0.1899	1.209	0.780 - 1.874	0.4
Dipstick Hematuria (%)				
+	0.0795	1.083	0.961 - 1.220	0.2
2+	0.1349	1.144	0.982 - 1.333	0.08
3+	0.4384	1.550	1.295 - 1.856	<0.001
Missing	0.0884	1.092	0.846 - 1.411	0.5
Model 2 (full model)				
Risk factor	Regression coefficient	odds ratios	95 percent confidence intervals	P value
Intercept	2.7879	-	-	<0.001
Age (years)	0.0279	1.028	1.024 - 1.033	<0.001
eGFR (ml/min/1.73m^2)	-0.0817	0.922	0.919 - 0.924	<0.001
Dipstick Proteinuria (%)				
+	1.1010	3.007	2.137 - 4.230	<0.001
Missing	0.1601	1.174	0.756 - 1.822	0.5
Dipstick Hematuria (%)				
+	0.0907	1.095	0.971 - 1.234	0.1
2+	0.1397	1.150	0.986 - 1.341	0.07
3+	0.4574	1.580	1.319 - 1.893	<0.001
Missing	0.0962	1.101	0.852 - 1.423	0.5
Body Mass Index (kg/m^2)	0.0067	1.007	0.995 - 1.018	0.2
Systolic blood pressure (mmHg)	0.0049	1.005	1.003 - 1.007	<0.001
Medicated for hypertension	0.1029	1.108	1.015 - 1.210	0.02
Glucose tolerance				
Fasting serum glucose (mg/dl)	-0.0030	0.997	0.995 - 0.999	<0.001
Non-fasting serum glucose (mg/dl)	-0.0034	0.997	0.995 - 0.998	<0.001

Medicated for diabetes mellitus	0.6224	1.863	1.453 - 2.390	<0.001
Smoking status				
Past smoker	-0.1243	0.883	0.520 - 1.500	0.6
Current smoker	0.3285	1.389	1.166 - 1.655	<0.001
Alcohol intake				
Occasional	-0.1964	0.822	0.706 - 0.957	0.01
Everyday	-0.1951	0.823	0.680 - 0.995	0.04

eGFR: estimated glomerular filtration rate at baseline

Table 4 Predictive equations for chronic kidney disease at 10 years	
Risk for CKD at 10 years=1/1+exp(-(intercept + sum of the estimate per risk factors))	
Model 1 (Simple model)	
Men	intercept + sum of the estimate per risk factors=3.3627 + 0.0404×age(years) -0.0932×eGFR(ml/min/1.73m^2)
	+1.5395(if proteinuria(+)) + 0.4237(if hematuria(+))
Women	intercept + sum of the estimate per risk factors=3.0796 + 0.0309×age(years) -0.0816×eGFR(ml/min/1.73m^2)
	+1.1677(if proteinuria(+)) + 0.4384(if hematuria(3+))
Model 2 (Full model[a])	
Men	intercept + sum of the estimate per risk factors=1.7047 + 0.0402×age(years) -0.0928×eGFR(ml/min/1.73m^2)
	+1.4492(if proteinuria(+)) + 0.4124(if hematuria(+))
	+0.0323×BMI(kg/m^2) + 0.0077×SBP(mmHg) + 0.2428 (if medicated for hypertension)
	-0.0017×fasting serum glucose(mg/dl) -0.0016×non-fasting serum glucose(mg/dl) + 0.2078(if medicated for diabetes mellitus)
	-0.1581(if past smoker) + 0.2313 (if current smoker) - 0.1639 (if daily drinker)
Women	intercept + sum of the estimate per risk factors=2.7879 + 0.0279×age(years) -0.0817×eGFR(ml/min/1.73m^2)
	+1.1010(if proteinuria(+)) + 0.4574(if hematuria(3+))
	+ 0.0049×SBP(mmHg) + 0.1029 (if medicated for hypertension)
	-0.0030×fasting serum glucose(mg/dl) -0.0034×non-fasting serum glucose(mg/dl) + 0.6224(if medicated for diabetes mellitus)
	+0.3285 (if current smoker) - 0.1964(if occasional drinker) - 0.1951 (if daily drinker)

eGFR: estimated glomerular filtration rate at baseline
BMI: body mass index
SBP: systolic blood pressure
[a] use fasting serum glucose or non-fasting serum glucose according to condition

Figure Legends

Figure 1 Calibration plots of predictive equations for CKD risk at 10 years among men

(A) Model 1, (B) Model 2. CKD, chronic kidney disease.

Figure 2 Calibration plots of predictive equations for CKD risk at 10 years among women

(A) Model 1, (B) Model 2. CKD, chronic kidney disease.

Supplemental Table 1 How points were assigned

Risk Factors	Regression coefficients	P value	Regression coefficients/constant[A]	Points awarded[B]
Intercept	-6.4049	<0.001		
Age (years)	0.0536	<0.001		
40-44[C]			0.00	0
45-49			1.00	1
50-54			2.00	2
55-59			3.00	3
60-64			4.00	4
65-69			5.00	5
70-74			6.00	6
eGFR (ml/min/1.73m^2)				
60-74	2.6370	<0.001	9.84	10
75-89	1.2163	<0.001	4.54	5
90-119	Reference			0
120-	-1.3529	0.001	-5.05	-5
Dipstick proteinurea				
+	1.5038	<0.001	5.61	6
Missing	-0.1391	0.6	-	0
Dipstick hematourea				
+	0.4126	<0.001	1.54	2
2+	0.2231	0.2	-	0
3+	0.3492	0.1	-	0

[A] constant = 0.0536 × 5 =0.268 (1 point was determined as the risk associated with a 5-year increase in age).

[B] Regression coefficients/constants were rounded to the nearest integer.

[C] People aged 40–44 were used for Reference.

Supplemental Table 3 Estimated risk for chronic kidney disease at 10 years stratified by sex

Risk Score (points)	Men		Women	
	Model 1	Model 2	Model 1	Model 2
-13				1%
-12				1%
-11				1%
-10				1%
-9				1%
-8		0%		1%
-7		0%	1%	1%
-6		0%	1%	1%
-5	0%	0%	1%	1%
-4	1%	1%	1%	1%
-3	1%	1%	2%	2%
-2	1%	1%	2%	2%
-1	1%	1%	2%	2%
0	2%	1%	2%	2%
1	2%	2%	2%	2%
2	3%	2%	3%	3%
3	3%	3%	3%	3%
4	4%	4%	4%	3%
5	6%	5%	4%	4%
6	7%	6%	4%	4%
7	9%	8%	5%	4%
8	12%	10%	6%	5%
9	15%	13%	6%	5%
10	19%	16%	7%	6%
11	23%	20%	8%	7%
12	28%	24%	9%	7%
13	34%	29%	10%	8%
14	40%	35%	11%	9%
15	47%	41%	12%	10%
16	53%	47%	13%	11%
17	60%	54%	15%	12%
18	66%	60%	17%	13%
19	72%	66%	18%	14%
20	77%	71%	20%	16%
21	81%	76%	22%	17%
22	85%	81%	24%	19%
23	88%	84%	27%	20%
24	91%	87%	29%	22%
25		90%	32%	24%
26		92%	34%	26%
27		94%	37%	28%
28		95%	40%	30%
29			43%	32%
30			46%	35%
31			49%	37%
32			52%	40%
33			55%	42%
34			58%	45%
35			61%	47%
36			64%	50%
37			66%	53%
38			69%	55%
39			72%	58%
40			74%	60%
41			76%	63%
42			78%	65%
43				68%
44				70%
45				72%
46				74%
47				76%
48				78%
49				80%
50				81%
51				83%
52				84%
53				86%
54				87%
55				88%

Supplemental Table 2 Risk for chronic kidney disease at 10 years according to risk category stratified by sex

			Points awarded for men	Points awarded for women
Model 1 (simple model)	Age (years)	40-44	0	0
		45-49	1	1
		50-54	2	2
		55-59	3	3
		60-64	4	4
		65-69	5	5
		70-74	6	6
	Dipstick Proteinuria	+	6	10
	Dipstick Hematuria	+	2	0
		2+	0	0
		3+	0	4
	eGFR (ml/min/1.73m^2)	60-74	10	22
		75-89	5	11
		90-119	0	0
		120-	-5	-7
Model 2 (full model)	Age (years)	40-44	0	0
		45-49	1	1
		50-54	2	2
		55-59	3	3
		60-64	4	4
		65-69	5	5
		70-74	6	6
	Dipstick Proteinuria	+	6	11
	Dipstick Hematuria	+	2	0
		2+	0	0
		3+	0	4
	eGFR (ml/min/1.73m^2)	60-74	10	25
		75-89	5	12
		90-119	0	0
		120-	-5	-8
	Blood pressure	Stage 1 hypertension[a]	1	1
		Stage 2 hypertension[b]	1	3
		Stage 3 hypertension[c]	3	0
		Medicated for hypertension	2	2
	Glucose tolerance	Impaired glucose tolerance[d]	-1	-3
		Diabetes[e]	0	0
		Medicated for diabetes	0	3
	Smoking	Past smoker	-1	0
		Current smoker	1	3
	Alcohol intake	Occasional	0	-2
		Everyday	-1	-2

eGFR: estimated glomerular filtration rate at baseline.

[a] 140 mmHg ≤ systolic blood pressure < 160 mmHg and/or 90 mmHg ≤ diastolic blood pressure < 100 mmHg.

[b] 160 mmHg ≤ systolic blood pressure < 180 mmHg and/or 100 mmHg ≤ diastolic blood pressure < 110 mmHg.

[c] 180 mmHg ≤ systolic blood pressure and/or 110 mmHg ≤ diastolic blood pressure.

[d] 110 mg/dL ≤ fasting serum glucose < 126 mg/dL or 140 mg/dL ≤ non-fasting serum glucose < 200 mg/dL.

[e] 126 mg/dL ≤ fasting serum glucose or 200 mg/dL ≤ non-fasting serum glucose.

Supplemental Table 4 Multivariable-adjusted odds ratios of chronic kidney disease (stage 3 or higher and/or proteinurea 2+ or 3+) among people living in the north of Ibaraki Prefecture, Japan

	Men					P value	Women					P value
N	17892						40963					
CKD event	2840						4660					
Model 1 (simple model)												
Age (years)	1.06	(	1.05	-	1.06	)	<0.001	1.02	(	1.02	-	1.03	)	<0.001
Dipstick proteinurea														
+	4.50	(	3.31	-	6.12	)	<0.001	3.20	(	2.29	-	4.46	)	<0.001
Missing	0.87	(	0.48	-	1.59	)	0.6	1.09	(	0.70	-	1.68	)	0.7
Dipstick hematourea														
+	1.51	(	1.22	-	1.88	)	<0.001	1.08	(	0.96	-	1.22	)	0.2
2+	1.25	(	0.88	-	1.77	)	0.2	1.13	(	0.97	-	1.32	)	0.1
3+	1.42	(	0.92	-	2.19	)	0.1	1.55	(	1.30	-	1.86	)	<0.001
Missing				-				1.31	(	1.02	-	1.70	)	0.04
eGFR (ml/min/1.73m^2)														
60-74	13.97	(	12.18	-	16.03	)	<0.001	14.64	(	13.17	-	16.26	)	<0.001
75-89	3.38	(	2.92	-	3.90	)	<0.001	3.58	(	3.21	-	3.99	)	<0.001
90-119				Reference							Reference			
120-	0.26	(	0.11	-	0.58	)	<0.001	0.43	(	0.31	-	0.59	)	<0.001
Model 2 (full model)														
Age (years)	1.05	(	1.05	-	1.06	)	<0.001	1.02	(	1.02	-	1.03	)	<0.001
Dipstick proteinurea														
+	4.34	(	3.17	-	5.92	)	<0.001	3.12	(	2.23	-	4.35	)	<0.001
Missing	0.87	(	0.48	-	1.59	)	0.7	1.05	(	0.68	-	1.63	)	0.8
Dipstick hematourea														
+	1.48	(	1.19	-	1.85	)	<0.001	1.09	(	0.97	-	1.23	)	0.2
2+	1.22	(	0.86	-	1.73	)	0.3	1.14	(	0.98	-	1.33	)	0.09
3+	1.37	(	0.88	-	2.12	)	0.2	1.58	(	1.32	-	1.88	)	<0.01
Missing				-				1.31	(	1.01	-	1.69	)	0.04
eGFR (ml/min/1.73m^2)														
60-74	14.05	(	12.22	-	16.14	)	<0.001	14.61	(	13.14	-	16.23	)	<0.001
75-89	3.42	(	2.96	-	3.96	)	<0.001	3.58	(	3.21	-	3.99	)	<0.001
90-119				Reference							Reference			
120-	0.26	(	0.12	-	0.59	)	0.001	0.43	(	0.31	-	0.58	)	<0.001
Blood pressure level														
Stage 1 hyp	1.24	(	1.11	-	1.39	)	<0.001	1.11	(	1.02	-	1.21	)	0.02
Stage 2 hyp	1.44	(	1.18	-	1.75	)	<0.001	1.40	(	1.18	-	1.66	)	<0.001
Stage 3 hyp	1.95	(	1.28	-	2.95	)	0.002	0.83	(	0.49	-	1.40	)	0.5
Medicated fo	1.62	(	1.44	-	1.83	)	<0.001	1.25	(	1.15	-	1.37	)	<0.001
Missing				-				<0.01	(	<0.01	-	>99.99	)	0.9
Glucose tolerance														
Impaired glu	0.78	(	0.68	-	0.89	)	<0.001	0.76	(	0.67	-	0.86	)	<0.001
Diabetes[c]	0.96	(	0.77	-	1.18	)	0.7	0.88	(	0.70	-	1.10	)	0.3
Medicated fo	1.09	(	0.85	-	1.40	)	0.5	1.41	(	1.12	-	1.77	)	<0.001
Missimg	0.40	(	0.03	-	5.95	)	0.5	0.69	(	0.15	-	3.24	)	0.6
Smoking status														
Past smoker	0.88	(	0.78	-	0.99	)	0.03	0.91	(	0.53	-	1.54	)	0.7
Current smok	1.23	(	1.09	-	1.37	)	<0.001	1.39	(	1.17	-	1.66	)	<0.001
Alcohol intake														
Occasional	0.99	(	0.86	-	1.14	)	0.9	0.83	(	0.71	-	0.97	)	0.02
Everyday	0.84	(	0.76	-	0.92	)	<0.001	0.83	(	0.69	-	1.00	)	0.06

eGFR: estimated glomerular filtration rate at baseline.

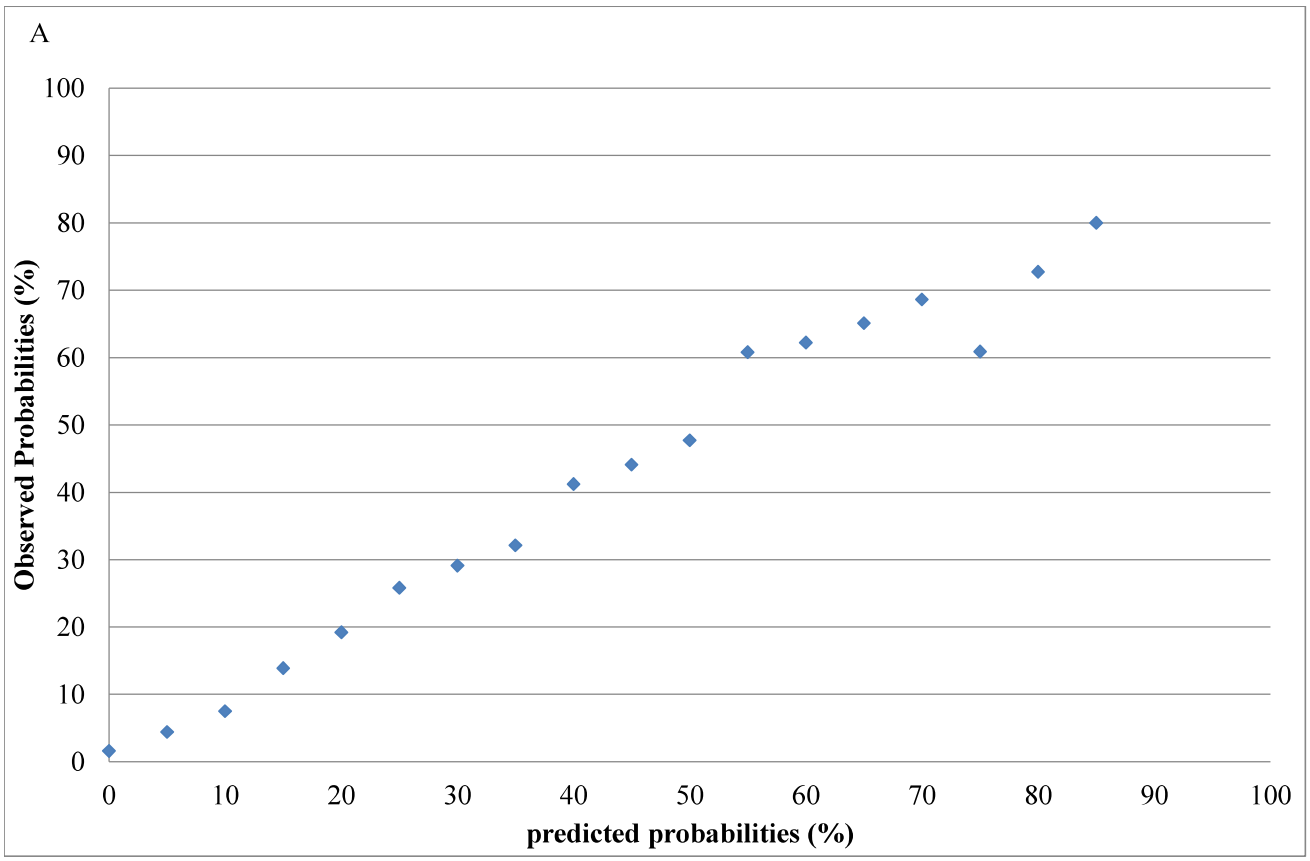
[a] 140 mmHg ≤ systolic blood pressure < 160 mmHg and/or 90 mmHg ≤ diastolic blood pressure < 100 mmHg.

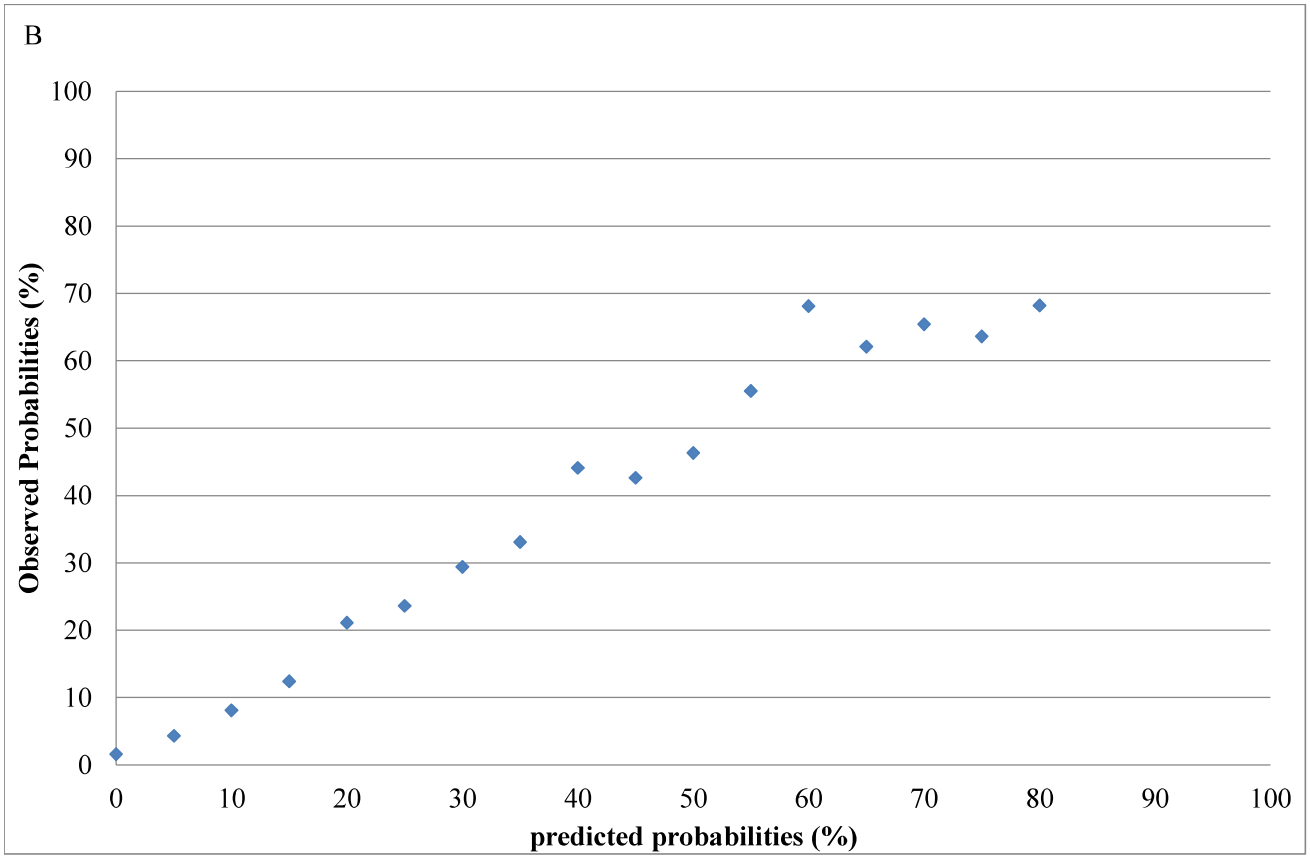
[b] 160 mmHg ≤ systolic blood pressure < 180 mmHg and/or 100 mmHg ≤ diastolic blood pressure < 110 mmHg.

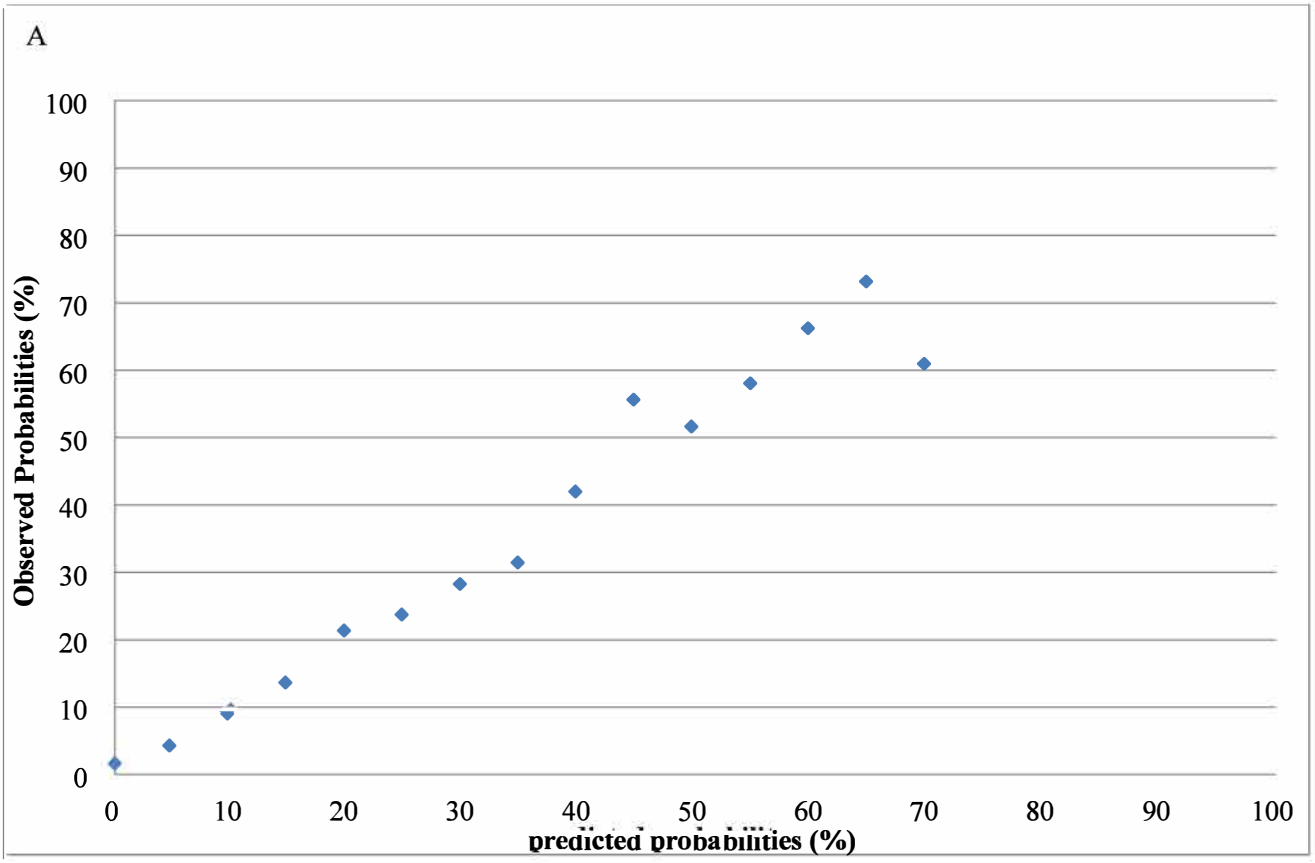
[c] 180 mmHg ≤ systolic blood pressure and/or 110 mmHg ≤ diastolic blood pressure.

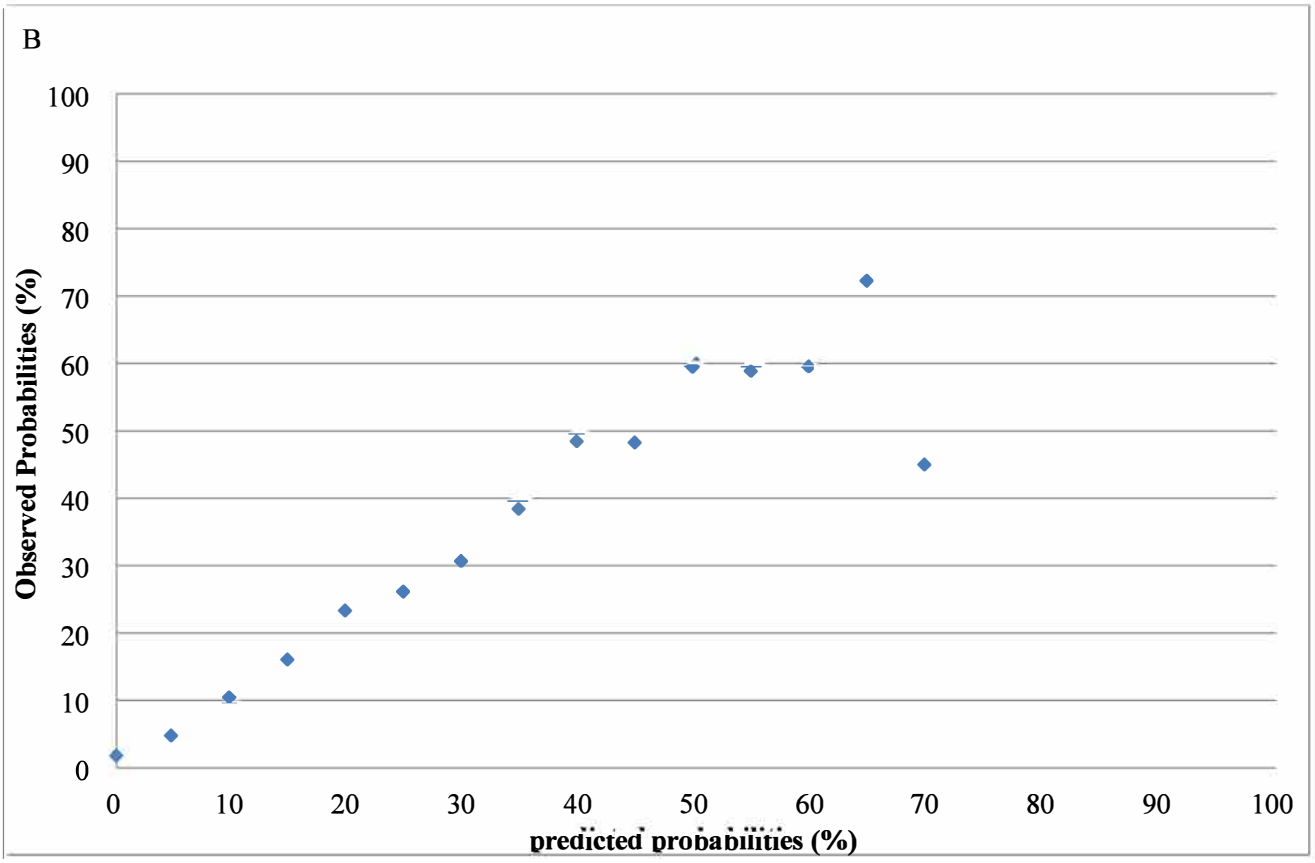
[d] 110 mg/dL ≤ fasting serum glucose < 126 mg/dL or 140 mg/dL ≤ non-fasting serum glucose < 200 mg/dL .

[e] 126 mg/dL ≤ fasting serum glucose or 200 mg/dL ≤ non-fasting serum glucose.


A
100
90
80
70
60
50
40
30
20
10
0
Observed Probabilities (%)
0
10
20
30
40
50
60
70
80
90
100
predicted probabilities (%)


B
100
90
80
70
60
50
40
30
20
10
0
Observed Probabilities (%)
0
10
20
30
40
50
60
70
80
90
100
predicted probabilities (%)


A
100
90
80
70
60
50
40
30
20
10
0
Observed Probabilities (%)
0
10
20
30
40
50
60
70
80
90
100
predicted probabilities (%)


B
100
90
80
70
60
50
40
30
20
10
0
Observed Probabilities (%)
0
10
20
30
40
50
60
70
80
90
100
predicted probabilities (%)