ORIGINAL PAPER

Nagoya J. Med. Sci. **87**. 220–236, 2025
doi:10.18999/nagjms.87.2.220

Utility of long-term systolic blood pressure variability for predicting the development of type 2 diabetes mellitus

Zean Song[1], Yuanying Li[1], Young-Jae Hong[1], Chifa Chiang[1,2], Masaaki Matsunaga[3], Yupeng He[3], Atsuhiko Ota[3], Koji Tamakoshi[4] and Hiroshi Yatsuya[1]

[1]*Department of Public Health and Health Systems, Nagoya University Graduate School of Medicine, Nagoya, Japan*
[2]*Department of Global and Community Health, Nagoya City University Graduate School of Nursing, Nagoya, Japan*
[3]*Department of Public Health, Fujita Health University School of Medicine, Toyoake, Japan*
[4]*Department of Nursing, Nagoya University Graduate School of Medicine, Nagoya, Japan*

ABSTRACT

Better identification of individuals at high risk for type 2 diabetes mellitus (T2DM) requires risk-prediction models incorporating novel predictors. Accordingly, this study aimed to evaluate the merits of including long-term systolic blood pressure variability (SBPV) in predicting T2DM incidence in a Japanese cohort of 3017 participants (2446 men, 571 women; age, 36–65 years) in 2007, who were followed up until March 2019. Consecutive SBP values, recorded between 2003 and 2007, were regressed annually for each participant. The slope and root-mean-square error of the regression line were calculated for each individual to represent SBPV. The significance of SBPV was examined by adding it to a multivariate Cox model incorporating age, sex, smoking status, regular exercise, family history of diabetes, body mass index, blood levels of triglycerides, high-density lipoprotein cholesterol, and fasting blood glucose. The c-index, net reclassification improvement (NRI), and integrated discrimination improvement (IDI) were used to compare the performance of the prediction models without (Model 1) and with (Model 2) SBPV. During the 9.8-year follow-up period, 135 participants developed T2DM. Although a statistically significant difference in c-index between Model 1 (0.785) and Model 2 (0.786) was not found, the NRI (8.312% [$p < 0.001$]) and IDI (0.700% [$p = 0.012$]) demonstrated that the performance of Model 2 improved compared with Model 1. In conclusion, results suggested that long-term SBPV slightly improved predictive utility for T2DM when added to a conventional prediction model. The study was registered at University Hospital Medical Information Network Clinical Trial registry (UMIN000052544, https://www.umin.ac.jp/).

Keywords: diabetes, systolic blood pressure variability, prediction model

Abbreviations:
BMI: body mass index
BP: blood pressure
BPV: blood pressure variability
FBG: fasting blood glucose
HDL-c: high-density lipoprotein cholesterol
HR: hazard ratio

Received: July 10, 2024; Accepted: September 13, 2024
Corresponding Author: Hiroshi Yatsuya, MD, PhD
Department of Public Health and Health Systems, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan
E-mail: h828@med.nagoya-u.ac.jp

IDI: integrated discrimination improvement
NRI: net reclassification improvement
RMSE: root-mean-square error
SBP: systolic blood pressure
SBPV: systolic blood pressure variability
TG: triglyceride
T2DM: type 2 diabetes mellitus

INTRODUCTION

Diabetes mellitus (DM) is the third-leading and most rapidly increasing contributor to the global burden of disease, as assessed by disability-adjusted life years among individuals 50–74 years of age in 2019.[1] The International Diabetes Federation has projected that the number of individuals diagnosed with DM will reach 643 million by 2030 and 783 million by 2045, posing an enormous challenge to individual well-being and public health.[2] To mitigate the burden of type 2 DM (T2DM) on society by reducing its incidence, it would be highly beneficial to develop preventive measures using risk prediction models to identify individuals with an elevated risk for T2DM.[3-5]

Long-term blood pressure variability (BPV), often referred to as "visit-to-visit BPV", emerges over the span of months or years between clinic visits. Research has linked BPV to various adverse health outcomes including stroke, cognitive dysfunction, and all-cause mortality.[6-8] We and others have recently reported an association between systolic BPV (SBPV) and the development of subsequent T2DM.[9,10] However, a statistically significant association does not necessarily indicate utility in prediction[11] beyond a model incorporating conventional factors.[12-14] As such, the present study aimed to investigate whether incorporating long-term SBPV data into an existing predictive model could improve prediction of the development of T2DM.

MATERIAL AND METHODS

Population

The Aichi Workers' Cohort Study is a prospective cohort study of non-communicable diseases including diabetes and cardiovascular diseases. It began in 1997 by recruiting employees from 2 worksites (a local government and a manufacturing company) in Aichi Prefecture, Japan. Because follow-up in the manufacturing company ended in 2002, the baseline year for the present analysis was set to 2007, which enabled the construction of a year–SBP regression line using up to five years' of SBP data before the start of the follow-up. Baseline data for this analysis were collected in 2007 from 3520 participants (2879 men, 641 women) between 36 and 65 years of age, with the majority engaged in clerical roles. All participants provided informed consent. Individuals with a history of diabetes at baseline (n = 176) and those lacking essential covariate information (n = 327) were excluded; ultimately, therefore, the analysis included data from 3017 participants. The study protocol was approved by the Ethics Review Committee of the Nagoya University Graduate School of Medicine, Nagoya, Japan (Approval number, 504-7). This study was registered with University Hospital Medical Information Network Clinical Trial registry (UMIN) Clinical Trials Registry (UMIN000052544, registered October 19, 2023).

Measurements

Participant weight was recorded to the nearest 0.5 kg, and height was measured to the nearest 0.1 cm. These measurements were performed with the participants wearing standard indoor attire without shoes. Body mass index (BMI) was calculated as weight in kilograms divided by height in meters squared (kg/m^2). Venous blood samples were collected from participants who had fasted for a minimum of 8 h or overnight. Serological tests were performed to analyze fasting blood glucose (FBG), high-density lipoprotein cholesterol (HDL-c), and triglyceride (TG) levels.

BP readings were obtained in accordance with the Japanese Society of Hypertension Guidelines for the Management of Hypertension 2000 (JSH 2000). BP measurements were performed after a 5 min rest period. Participants were instructed to refrain from consuming coffee or tobacco for at least 30 min before the examination. The majority of BP measurements were acquired using automated oscillometric BP monitoring devices (specifically, BP-103i II and BP-103N II, Nippon Colin Co, Ltd, Tokyo, Japan) while participants were in a seated position. Typically, only 1 BP measurement per individual was recorded, unless SBP was ≥140 mmHg and/or diastolic BP was ≥90 mmHg, in which case an additional measurement was performed, with the lower of the 2 values recorded.

Definition of risk factors

Age was considered to be a continuous variable. BMI was classified into 4 categories: <23; 23 to <25 (reference); 25 to <27.5; and ≥27.5 kg/m^2. Obesity was defined[15] as BMI ≥25 kg/m^2. Smoking status was dichotomized as non-current (reference) versus (vs) current. Regular exercise was defined as 60 min/day for ≥12 days/month. Alcohol consumption was estimated from the frequency of drinking and the amount of alcohol per occasion, which was divided into 4 groups: 0 (reference); <23; 23 to <46; and ≥46 g/day. A family history of diabetes was defined as the presence of first-degree relatives(s) (parents, sibling[s], or children) with the disease. History of dyslipidemia was dichotomized as self-reported antihyperlipidemic medication use or physician-diagnosed hyperlipidemia. Blood levels of TG and HDL-c were categorized into 2 groups: ≥40 mg/dL (reference) or <40 mg/dL; and <150 (reference) or ≥150 mg/dL, respectively. FBG levels were grouped into 3 categories: <100 (reference); 100–110; and 110–126 mg/dL.

Baseline SBP was treated as a continuous variable. SBPV was assessed using 2 indices: SBP change trend; and root-mean-square error (RMSE). The SBP change trend, representing the slope of the year–SBP regression line from 2003 to 2007, was considered to be a continuous variable. The RMSE of SBP, which is the standard deviation of SBP around the year–SBP regression line, was divided into tertiles, in ascending order, as follows: T1 (reference), T2, and T3.[16]

Follow-up and ascertainment of incident T2DM

Participants were monitored until the occurrence of T2DM, cessation of employment, death, or until March 31, 2019, whichever occurred first. However, retirees who consented to participate via mail were included in the cohort. In cases of death during employment, confirmation was obtained through the worksite's healthcare division or via notification by the next of kin during biennial postal follow-up surveys.

Cases of incident T2DM were identified through compulsory annual health check-ups conducted at the workplace until retirement, as well as through questionnaire surveys administered during employment and post-retirement. The former was defined as an FBG level ≥126 mg/dL and, if available, a glycated hemoglobin (HbA1c) level ≥6.5% (in accordance with the United States National Glycohemoglobin Standardization Program method). However, HbA1c testing was only performed in patients with positive urinary glucose levels. A self-administered questionnaire survey was conducted approximately every 2 years from 2006 to 2019. Participants were requested

to report any medical history related to specific conditions, including T2DM. Those who reported a history of T2DM were further asked to provide information about the physician overseeing their disease management so that medical records could be verified.

Statistical analysis

A univariate Cox hazard model was used to determine hazard ratio (HR) and corresponding 95% confidence interval (CI) for each predictor. A multivariable Cox proportional hazard model that forcedly entered variables, including sex, age, BMI, smoking status, alcohol consumption, regular exercise, family history of diabetes, history of dyslipidemia, baseline SBP, FBG, TG, and HDL-c, was used as a base model (Model 1). Model 2 included the variables from Model 1 plus long-term SBPV. A backward elimination method, using $p < 0.1$, was used to confirm whether the SBPV variable remained in the model.

The prediction models were assessed for discrimination using c-index, which can have values from 0.5 (no discriminative ability) to 1.0 (perfect discrimination), with values <0.7, 0.7 to <0.8, and ≥0.8 considered to be poor, acceptable, and excellent discrimination, respectively.[17] The 2 c-indexes were compared by calculating the z-value to obtain a *p*-value (equation not shown).[18] For the assessment of calibration—more specifically, how closely the predicted risks fit the actual risks—a calibration plot was used. Kaplan–Meier survival analysis was used to determine the T2DM event rate, which was then plotted against the event rate estimated by the prediction model. To draw a calibration plot, participants were categorized into 4 groups according to quartile of the predicted rate.[19] The slope and intercept of the regression line were used as calibration measures. The calibration slope indicates the extent to which the estimated risk matches the observed risk. Thus, the target value of the slope was 1, and that of the intercept was 0.[20] The model was also calibrated using Greenwood-Nam-D'Agostino test, which is an extension of Hosmer–Lemeshow test for situations with censored survival data.[21]

Improvement in model performance was evaluated using categorical net reclassification improvement (NRI) and integrated discrimination improvement (IDI).[22] The NRI quantified the appropriateness of the change in the predicted probabilities of T2DM between the 2 models. Reclassification tables for participants who did (event group) and did not (nonevent group) develop T2DM were constructed using the 10-year risk categories <0.05, 0.05 to <0.10, and ≥0.10.[23] The NRIs for the event and nonevent groups were calculated separately to yield the overall NRI. A z-score was calculated to determine whether the difference between the 2 models was statistically significant.[18] The IDI is another—and, reportedly more sensitive—summary statistic of the area under the receiver operating characteristic curve (AUC) to evaluate the incremental value of an added predictor.[24] The IDI can also be interpreted as the difference in the discrimination slopes between Models 1 and 2.[25] This was calculated based on the probability predicted for each participant using the 2 models. The z-score was calculated to determine whether the differences between the 2 models were statistically significant.[18]

Several sensitivity analyses were performed, the first of which excluded participants who reported antihypertensive medication use in any year from 2003 to 2007 (n = 574) because the initiation of or adherence to antihypertensive therapy could be a source of long-term BPV. Second, an analysis that censored participants at 60 years of age was performed due to the dependence on self-reports for the ascertainment of T2DM incidence after retirement, by which annual mandatory health checkups were also conducted by the employer. Third, sex-stratified analysis was performed. Differences with $p < 0.05$ were considered to be statistically significant, and all analyses were performed using R version 4.2.1 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

The mean age of the cohort at baseline was 48.7 years and the proportion of male participants was 81.1%. The prevalence of obesity (ie, BMI ≥25 kg/m^2), current smoker, and alcohol consumption ≥46 g/day were 19.8%, 21.7%, and 2.8%, respectively. Twenty-one percent of participants engaged in regular exercise, 3.1% had a family history of diabetes and 5.6% had a history of dyslipidemia. The mean baseline SBP was 124 mmHg. Of the participants, 3.8% had FBG values of 110–126 mmHg; 21.5% had TG values ≥150 mg/dL; and 5.1% had HDL-c levels <40 mg/dL. The mean SBP change trend in the participants was 0.19 mmHg/year (Table 1). Thirty-five participants developed T2DM during a median follow-up of 9.8 years.

In the univariate models, sex, age, BMI, smoking status, alcohol consumption, history of dyslipidemia, baseline SBP, FBG, TG, and HDL-c, SBP change slope, and SBP RMSE were significantly associated with the incidence of T2DM. The multivariate Cox hazard model with a backward selection procedure confirmed that the SBP change slope and RMSE remained in the prediction model (Supplementary Table 1).

Table 1 Characteristics of participants and univariable HR (95% CI) for type 2 diabetes incidence, Aichi, Japan, 2007–2019

Risk factors	Mean (SD) or N (%)	Univariable HR (95% CI)
Age, year	48.7 (6.8)	1.04 (1.01–1.07)
Sex		
Female	571 (18.9)	1 (Reference)
Male	2446 (81.1)	1.95 (1.12–3.40)
Body mass index, kg/m^2		
<23.0	1654 (54.8)	1 (Reference)
23.0–<25.0	765 (25.4)	2.19 (1.42–3.38)
25.0–<27.5	429 (14.2)	2.94 (1.84–4.69)
≧27.5	169 (5.6)	5.47 (3.26–9.19)
Smoking status		
Non-current	2361 (78.3)	1 (Reference)
Current	656 (21.7)	1.49 (1.03–2.17)
Alcohol consumption, g/day		
0	791 (26.2)	1 (Reference)
<23	1792 (59.4)	0.86 (0.58–1.27)
23–<46	351 (11.6)	0.93 (0.51–1.70)
≧46	83 (2.8)	2.15 (1.00–4.63)
Regular exercise		
Yes	633 (21.0)	1 (Reference)
No	2384 (79.0)	0.98 (0.65–1.47)
Family history of diabetes		
No	2922 (96.9)	1 (Refence)
Yes	95 (3.1)	1.23 (0.50–3.00)
History of dyslipidemia		
No	2849 (94.4)	1 (Reference)
Yes	168 (5.6)	2.14 (1.23–3.72)

Baseline SBP, mmHg	124 (15)	1.03 (1.02–1.04)
Fasting blood glucose, mg/dL		
<100	2411 (79.9)	1 (Reference)
100–<110	492 (16.3)	4.20 (2.89–6.09)
110–<126	114 (3.8)	9.77 (6.00–15.89)
Triglycerides, mg/dL		
<150	2367 (78.5)	1 (Reference)
≧150	650 (21.5)	2.19 (1.54–3.11)
High-density lipoprotein cholesterol, mg/dL		
≧40	2864 (94.9)	1 (Reference)
<40	153 (5.1)	3.06 (1.88–4.98)
SBP change trend, mmHg/year	0.19 (4.7)	1.04 (1.01–1.08)
SBP RMSE, mmHg		
T1	1012 (33.5)	1 (Reference)
T2	999 (33.1)	1.49 (0.94–2.34)
T3	1006 (33.3)	1.99 (1.29–3.07)

HR: hazard ratio
CI: confidence interval
N: number
SD: standard deviation
SBP: systolic blood pressure
RMSE: root-mean-square error

In Model 1, higher BMI (≥27.5 kg/m^2, HR 3.62 [95% CI 2.09–6.28]), higher baseline SBP (HR 1.01 [95% CI 1.00–1.02), higher FBG (110 to <126 mg/dL, HR 8.13 [95% CI 4.82–13.7]), and lower HDL-c (HR 1.98 [95% CI 1.15–3.41]) were significantly associated with a higher incidence of T2DM. Model 2 results revealed that higher BMI (≥27.5 kg/m^2, HR 3.51 [95% CI 2.02–6.09]), higher FBG (110 to <126 mg/dL, HR 8.04 [95% CI 4.78–13.54]), lower HDL-c (HR 1.93 [95% CI 1.12–3.33]), and SBP RMSE tertile 3 (HR 1.77 [95% CI 1.13–2.77]) were significantly and positively associated with higher risks for developing T2DM (Table 2).

Table 2 Multivariable Cox regression HR (95% CI) of the type 2 diabetes mellitus risk prediction model, Aichi, Japan, 2007–2019

Risk factors	Model 1 HR (95% CI)	Model 2 HR (95% CI)
Age, year	1.01 (0.98–1.04)	1.01 (0.98–1.04)
Sex		
Female	1 (reference)	1 (reference)
Male	1.12 (0.62–2.04)	1.22 (0.67–2.24)
Body mass index, kg/m^2		
<23.0	1 (reference)	1 (reference)
23.0–<25.0	1.55 (0.99–2.43)	1.56 (0.99–2.43)
25.0–<27.5	2.08 (1.28–3.39)	2.05 (1.26–3.34)
≧27.5	3.62 (2.09–6.28)	3.51 (2.02–6.09)

Smoking status		
Non–current	1 (reference)	1 (reference)
Current	1.26 (0.84–1.87)	1.23 (0.83–1.84)
Alcohol consumption, g/day		
0	1 (reference)	1 (reference)
<23	0.79 (0.52–1.18)	0.81 (0.54–1.22)
23–<46	0.68 (0.36–1.23)	0.71 (0.38–1.34)
≧46	1.52 (0.69–3.39)	1.59 (0.72–3.55)
Regular exercise		
Yes	1 (reference)	1 (reference)
No	0.97 (0.64–1.47)	0.99 (0.65–1.50)
Family history of diabetes		
No	1 (reference)	1 (reference)
Yes	1.10 (0.45–2.72)	1.15 (0.47–2.85)
History of dyslipidemia		
No	1 (reference)	1 (reference)
Yes	1.18 (0.67–2.10)	1.25 (0.70–2.22)
Baseline SBP, mmHg	1.01 (1.00–1.02)	1.00 (0.99–1.02)
Fasting blood glucose, mg/dL		
<100	1 (reference)	1 (reference)
100–<110	3.48 (2.37–5.12)	3.51 (2.39–5.17)
110–<126	8.13 (4.82–13.71)	8.04 (4.78–13.54)
Triglycerides, mg/dL		
<150	1 (reference)	1 (reference)
≧150	1.27 (0.86–1.87)	1.28 (0.87–1.89)
High–density lipoprotein cholesterol, mg/dL		
≧40	1 (reference)	1 (reference)
<40	1.98 (1.15–3.41)	1.93 (1.12–3.33)
SBP change trend, mmHg/year	–	1.02 (0.98–1.06)
SBP RMSE, mmHg		
T1	–	1 (reference)
T2	–	1.29 (0.81–2.09)
T3	–	1.77 (1.13–2.77)

HR: hazard ratio
CI: confidence interval
SBP: systolic blood pressure
RMSE: root-mean-square error

Chi-squared values for the Greenwood-Nam-D'Agostino test for Model 1 and Model 2 were 4.07 (p = 0.25) and 6.94 (p = 0.07), respectively (Table 3). The calibration slopes and intercepts for Model 2 were 1.083 and -0.079 (Figure 1), which confirmed that the model was well-calibrated.

Table 3 Comparisons of the prediction models with or without SBP RMSE, Aichi, Japan, 2007–2019

	Model 1	Model 2	*p*
c-index (standard error)	0.785 (0.020)	0.786 (0.022)	0.386
Greenwood-Nam-D'Agostino test (p-values)	4.07 (0.25)	6.94 (0.07)	
NRI_event[†], %	8.000		
NRI_nonevent[‡], %	0.312		
NRI_overall[§], %	8.312		<0.001
IDI, %	0.700		0.018

SBP: systolic blood pressure
RMSE: root-mean-square error
NRI: net reclassification improvement
IDI: integrated discrimination improvement
[†]NRI_event indicates the NRI calculated for event group.
[‡]NRI_nonevent indicates the NRI calculated for nonevent group.
[§]NRI_overall is the sum of NRI_event and NRI_nonevent indicates the total NRI for model 2 compared with Model 1.
Model 1 and Model 2 predictors are shown in Table 2.

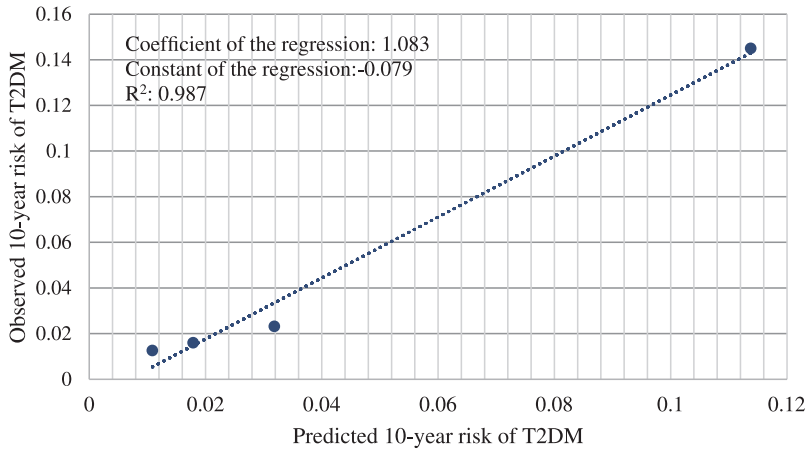

Fig. 1 Calibration plot for type 2 diabetes mellitus cases, showing predicted and observed 10-year risk according to quartiles of Model 2 predicted 10-year risk
T2DM: type 2 diabetes mellitus

The c-indices for Models 1 and 2 were 0.785 (standard error = 0.020) and 0.786 (standard error = 0.022), respectively; the difference between the two c-indices, however, was not statistically significant (p = 0.386). The overall NRI was 0.083, with a positive value indicating that the addition of SBPV improved the prediction, and the degree of improvement was statistically significant (p < 0.001). More specifically, event and non-event reclassification table analyses indicated that the event NRI was 0.080, and the non-event NRI was 0.003 (Supplementary Tables 2A and 2B). The IDI score was 0.007 (p = 0.018), indicating a significant improvement in the overall discriminative ability of the prediction model.

An analysis that excluded participants who reported antihypertensive medication use in any year from 2003 to 2007 found that the c-indices of Models 1 and 2 were 0.798 and 0.800, respectively. The overall NRI and IDI values were 9.243% ($p < 0.001$) and 1.3% ($p = 0.020$), respectively (Supplementary Table 3). These results demonstrated the robustness of the findings. Second, an analysis of censored participants 60 years of age revealed that the discrimination between Models 1 and 2 was 0.817 ($p = 0.023$) and 0.818 ($p = 0.023$), respectively (Supplementary Table 4). The analysis restricted to male participants yielded similar results (Supplementary Tables 5 and 6).

DISCUSSION

This Japanese cohort study examined the value of long-term SBPV in predicting 10-year T2DM risk when added to a conventional prediction model. The addition of SBPV significantly improved predictive ability in terms of discrimination and calibration.

To our knowledge, this is the first study to explore the potential of SBPV as a predictor of T2DM incidence.[26,27] Noble et al[28] conducted a systematic review of 94 risk models to describe their statistical properties and clinical applications of published models. The previous risk models included similar components, such as age, sex, smoking, BMI, physical activity, SBP, family history of diabetes, FBG, TG, HDL-c, and had similar discriminatory properties, with AUCs ranging from 0.74 to 0.85.[13,29-34] In the present study, we evaluated model performance using Harrell's C (ie, c-index)[18] and obtained a model that yielded a c-index of 0.786. Although the c-index did not increase (0.785 vs 0.786), the NRI and IDI revealed some improvement when SBPV was added as a predictor.[35] This improvement meant that SBPV aided in the identification of individuals who would develop diabetes, thus enabling us to efficiently deliver preventive services.[36] IDI, which would reflect the difference in the discrimination curves between 2 models, also indicated some improvement, although we caution that it could provide spurious results.[37] Collectively, our results suggest that the inclusion of SBPV as a predictor may slightly improve predictive capability.

However, limitations of the present study warrant further consideration. First, we were unable to conclude whether the model can be applied to women. Future studies should be conducted in different populations predominantly consisting of women. Second, although the present study found that SBPV could improve predictive ability, it is difficult to determine whether the improvement was significant from a preventive medicine or clinical perspective, because few studies have described the application and use of prediction models as interventions to change outcomes.[38-40] Further research is required to examine the significance and utility of SBPV. Third, the single measurement of FBG and HbA1c levels used in the present study to define T2DM does not officially meet the diagnostic criteria of the Japan Diabetes Society.[41] In addition, accurate classification of T2DM requires the measurement of islet-associated autoantibodies, such as glutamic acid decarboxylase.[42] The same approach has been widely used in previous epidemiological studies in Japan[43,44] and elsewhere.[45,46] The T2DM cases in the present study may have included non-diabetic cases or those with slowly progressive insulin-dependent diabetes mellitus. Future prospective studies using larger databases, especially electronic medical records containing accurate T2DM diagnoses, should address this issue.[47] Finally, even if additional BPV indicators are determined to be useful, more research investigating the pathogenesis of SBPV to determine reference threshold values of SBPV is required.

In conclusion, incorporating SBPV into a conventional model with predictors of BMI, TG, HDL-c, and FBG levels yielded a modest improvement in the predictive ability of the 10-year

T2DM risk model. Further research is necessary to explore the predictive power of SBPV for T2DM in different population settings as well as its practical application in preventive medicine.

ACKNOWLEDGEMENTS

Author contributions

HY and ZS conceptualized and designed the study, performed the analyses, drafted the initial manuscript, and revised the manuscript. HY, YL, and KT designed the study and collected data. All the other coauthors critically reviewed and revised the manuscript. All authors approved the final manuscript submitted for publication and agreed to be accountable for all aspects of this work.

Conflict of interest disclosure

There are no conflicts of interest to declare.

Financial disclosures

This work was supported in part by Ministry of Education, Culture, Sports, Science and Technology/Japan Society for the Promotion of Science (MEXT/JSPS) KAKENHI (grant numbers 09470112, 13470087 and 17390185 to Hideaki Toyoshima; 13770192, 15689011, 17790384, 20790438, 22390133, 23659346, 26293153 18H03057, 22H03349 and 23K24607 to HY; 12670352, 16590499, 18590594, 20590641, 23590787, 15K08802 and 20K10496 to KT; 25893088, 16K19278 and 19K19419 to YL); Health and Labor Sciences research grants for Comprehensive Research on Cardiovascular and Life-Style Related Diseases (H26-Junkankitou [Seisaku]-Ippan-001, H29-Junkankitou [Seishuu]-Ippan-003, 20FA1002, and 23FA1006) and for Research on Occupational Safety and Health (23JA1006) from the Ministry of Health, Labour and Welfare; and research grants from the Japan Atherosclerosis Prevention Fund, the Uehara Memorial Fund, and funds from the Noguchi Medical Research Institute, the Handashi Ishikai, General Inc Association, and the Aichi Health Promotion Foundation.

Meeting presentation

The authors acknowledge that portions of this manuscript were previously presented at the American Heart Association EPILifestyle 2023 Scientific Sessions, held between February 28 and March 3, 2023, in Boston, Massachusetts, USA. They presented a preliminary version of the findings titled “Predictive Value of Long-term Systolic Blood Pressure Variability for the Development of Type 2 Diabetes Mellitus”. The authors are grateful to the conference organizers for providing a platform for discussion and to the attendees for their valuable feedback and insightful questions, which have significantly contributed to the development of this manuscript.

Additional contributions

The authors are grateful to the participants as well as the staff of the health care division of the worksite. They also thank the technical assistance of Ms. Miho Ueda and Ms. Miho Takeuchi, and Professor Emeritus Hideaki Toyoshima who established the Aichi Workers’ Cohort Study.

REFERENCES

1 GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*.

2020;396(10258):1204–1222. doi:10.1016/S0140-6736(20)30925-9
2 Bouchi R, Noda M, Hayashino Y, et al. Behavioral change stage might moderate the impact of multifaceted interventions on non-attendance from medical care among patients with type 2 diabetes: The Japan Diabetes Outcome Intervention Trial-2 Large-Scale Trial 007 (J-DOIT2-LT007). *J Diabetes Investig.* 2023;14(7):907–916. doi:10.1111/jdi.14012
3 Zou Q, Qu K, Luo Y, Yin D, Ju Y, Tang H. Predicting Diabetes Mellitus With Machine Learning Techniques. *Front Genet.* 2018;9:515. doi:10.3389/fgene.2018.00515
4 Li Y, Yang M, Yuan L, Li T, Zhong X, Guo Y. Associations between a polygenic risk score and the risk of gestational diabetes mellitus in a Chinese population: a case-control study. *Endocr J.* 2023;70(12):1159–1168. doi:10.1507/endocrj.EJ23-0245
5 Miyamori D, Tanaka M, Furuhashi M, et al. Prediction of new onset of diabetes mellitus during a 10-year period by using a combination of levels of alanine aminotransferase and γ-glutamyl transferase. *Endocr J*. 2021;68(12):1391–1402. doi:10.1507/endocrj.EJ20-0823
6 Wang H, Li M, Xie SH, et al. Visit-to-visit Systolic Blood Pressure Variability and Stroke Risk: A Systematic Review and Meta-analysis. *Curr Med Sci.* 2019;39(5):741–747. doi:10.1007/s11596-019-2100-9
7 Chiu TJ, Yeh JT, Huang CJ, et al. Blood pressure variability and cognitive dysfunction: A systematic review and meta-analysis of longitudinal cohort studies. *J Clin Hypertens (Greenwich).* 2021;23(8):1463–1482. doi:10.1111/jch.14310
8 de Havenon A, Delic A, Yaghi S, et al. Midlife Blood Pressure Variability and Risk of All-Cause Mortality and Cardiovascular Events During Extended Follow-up. *Am J Hypertens.* 2021;34(12):1269–1275. doi:10.1093/ajh/hpab106
9 Zhou R, Li FR, Liu K, et al. Long-Term Visit-To-Visit Blood Pressure Variability and Risk of Diabetes Mellitus in Chinese Population: A Retrospective Population-Based Study. *Int J Public Health.* 2023;68:1605445. doi:10.3389/ijph.2023.1605445
10 Song Z, He Y, Chiang C, et al. Long-term variability and change trend of systolic blood pressure and risk of type 2 diabetes mellitus in middle-aged Japanese individuals: findings of the Aichi Workers' Cohort Study. *Hypertens Res.* 2022;45(11):1772–1780. doi:10.1038/s41440-022-00993-2
11 Folsom AR. Classical and novel biomarkers for cardiovascular risk prediction in the United States. *J Epidemiol*. 2013;23(3):158–162. doi:10.2188/jea.JE20120157
12 Mao Y, Zhu Z, Pan S, et al. Value of machine learning algorithms for predicting diabetes risk: A subset analysis from a real-world retrospective cohort study. *J Diabetes Investig.* 2023;14(2):309–320. doi:10.1111/jdi.13937
13 Chen L, Magliano DJ, Balkau B, et al. AUSDRISK: an Australian Type 2 Diabetes Risk Assessment Tool based on demographic, lifestyle and simple anthropometric measures. *Med J Aust.* 2010;192(4):197–202. doi:10.5694/j.1326-5377.2010.tb03507.x
14 Yatsuya H, Li Y, Hirakawa Y, et al. A Point System for Predicting 10-Year Risk of Developing Type 2 Diabetes Mellitus in Japanese Men: Aichi Workers' Cohort Study. *J Epidemiol.* 2018;28(8):347–352. doi:10.2188/jea.JE20170048
15 Ogawa W, Hirota Y, Miyazaki S, et al. Definition, criteria, and core concepts of guidelines for the management of obesity disease in Japan. *Endocr J.* 2024;71(3):223–231. doi:10.1507/endocrj.EJ23-0593
16 Yano Y. Visit-to-visit blood pressure variability—what is the current challenge? *Am J Hypertens*. 2017;30(2):112–114. doi:10.1093/ajh/hpw124
17 Lloyd-Jones DM. Cardiovascular risk prediction: basic concepts, current status, and future directions. *Circulation.* 2010;121(15):1768–1777. doi:10.1161/CIRCULATIONAHA.109.849166
18 Zhou ZR, Wang WW, Li Y, et al. In-depth mining of clinical data: the construction of clinical prediction model with R. *Ann Transl Med.* 2019;7(23):796. doi:10.21037/atm.2019.08.63
19 Crowson CS, Atkinson EJ, Therneau TM. Assessing calibration of prognostic risk scores. *Stat Methods Med Res.* 2016;25(4):1692–1706. doi:10.1177/0962280213497434
20 Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles heel of predictive analytics. *BMC Med.* 2019;17(1):230. doi:10.1186/s12916-019-1466-7
21 D'Agostino RB, Nam BH. Evaluation of the Performance of Survival Analysis Models: Discrimination and Calibration Measures. In: Balakrishnan N, Rao CR, eds. *Advances in Survival Analysis*. Elsevier; 2003:1–25. *Handbook of Statistics*; vol 23. https://doi.org/10.1016/S0169-7161(03)23001-7
22 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(2):157–172;discussion 207–212. doi:10.1002/sim.2929
23 Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations

to measure usefulness of new biomarkers. *Stat Med*. 2011;30(1):11–21. doi:10.1002/sim.4085
24 Bragg F, Trichia E, Aguilar-Ramirez D, Bešević J, Lewington S, Emberson J. Predictive value of circulating NMR metabolic biomarkers for type 2 diabetes risk in the UK Biobank study. *BMC Med.* 2022;20(1):159. doi:10.1186/s12916-022-02354-9
25 Pencina MJ, D'Agostino RB Sr, Demler OV. Novel metrics for evaluating improvement in discrimination: net reclassification and integrated discrimination improvement for normal variables and nested models. *Stat Med.* 2012;31(2):101–113. doi:10.1002/sim.4348
26 Choi BG, Rha SW, Kim SW, Kang JH, Park JY, Noh YK. Machine Learning for the Prediction of New-Onset Diabetes Mellitus during 5-Year Follow-up in Non-Diabetic Patients with Cardiovascular Risks. *Yonsei Med J.* 2019;60(2):191–199. doi:10.3349/ymj.2019.60.2.191
27 Joshi RD, Dhakal CK. Predicting Type 2 Diabetes Using Logistic Regression and Machine Learning Approaches. *Int J Environ Res Public Health.* 2021;18(14):7346. doi:10.3390/ijerph18147346
28 Noble D, Mathur R, Dent T, Meads C, Greenhalgh T. Risk models and scores for type 2 diabetes: systematic review. *BMJ.* 2011;343:d7163. doi:10.1136/bmj.d7163
29 Schmidt MI, Duncan BB, Bang H, et al. Identifying individuals at high risk for diabetes: The Atherosclerosis Risk in Communities study. *Diabetes Care.* 2005;28(8):2013–2018. doi:10.2337/diacare.28.8.2013
30 Rahman M, Simmons RK, Harding AH, Wareham NJ, Griffin SJ. A simple risk score identifies individuals at high risk of developing Type 2 diabetes: a prospective cohort study. *Fam Pract.* 2008;25(3):191–196. doi:10.1093/fampra/cmn024
31 Lindström J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. *Diabetes Care.* 2003;26(3):725–731. doi:10.2337/diacare.26.3.725
32 Wilson PW, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB Sr. Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. *Arch Intern Med.* 2007;167(10):1068–1074. doi:10.1001/archinte.167.10.1068
33 Stern MP, Williams K, Haffner SM. Identification of persons at high risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? *Ann Intern Med.* 2002;136(8):575–581. doi:10.7326/0003-4819-136-8-200204160-00006
34 Hippisley-Cox J, Coupland C, Robson J, Sheikh A, Brindle P. Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. *BMJ.* 2009;338:b880. doi:10.1136/bmj.b880
35 Thomas LE, O'Brien EC, Piccini JP, D'Agostino RB, Pencina MJ. Application of net reclassification index to non-nested and point-based risk prediction models: a review. *Eur Heart J.* 2019;40(23):1880–1887. doi:10.1093/eurheartj/ehy345
36 Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. *Ann Intern Med.* 2014;160(2):122–131. doi:10.7326/M13-1522
37 Hilden J, Gerds TA. A note on the evaluation of novel biomarkers: do not rely on integrated discrimination improvement and net reclassification index. *Stat Med.* 2014;33(19):3405–3414. doi:10.1002/sim.5804
38 Absetz P, Oldenburg B, Hankonen N, et al. Type 2 diabetes prevention in the real world: three-year results of the GOAL lifestyle implementation trial. *Diabetes Care.* 2009;32(8):1418–1420. doi:10.2337/dc09-0039
39 Kulzer B, Hermanns N, Gorges D, Schwarz P, Haak T. Prevention of diabetes self-management program (PREDIAS): effects on weight, metabolic risk factors, and behavioral outcomes. *Diabetes Care.* 2009;32(7):1143–1146. doi:10.2337/dc08-2141
40 Laatikainen T, Dunbar JA, Chapman A, et al. Prevention of type 2 diabetes by lifestyle intervention in an Australian primary health care setting: Greater Green Triangle (GGT) Diabetes Prevention Project. *BMC Public Health.* 2007;7:249. doi:10.1186/1471-2458-7-249
41 Seino Y, Nanjo K, Tajima N, et al. Report of the committee on the classification and diagnostic criteria of diabetes mellitus. *J Diabetes Investig*. 2010;1(5):212–228. doi:10.1111/j.2040-1124.2010.00074.x
42 Shimada A, Kawasaki E, Abiru N, et al. New diagnostic criteria (2023) for slowly progressive type 1 diabetes (SPIDDM): Report from Committee on Type 1 Diabetes in Japan Diabetes Society (English version). *Diabetol Int*. 2024;15(1):1–4. doi:10.1007/s13340-023-00679-1
43 Xu J, Goto A, Konishi M, et al. Development and validation of prediction models for the 5-year risk of type 2 diabetes in a Japanese population: Japan Public Health Center-based Prospective (JPHC) Diabetes Study. *J Epidemiol*. 2024;34(4):170–179. doi:10.2188/jea.JE20220329
44 Ikeda I, Fujihara K, Osawa T, et al. Retrospective cohort study to examine the association between serum amylase and the incidence of type 2 diabetes mellitus, Toranomon Hospital Health Management Center Study 23 (TOPICS 23). *BMJ Open Diabetes Res Care*. 2023;11(4):e003482. doi:10.1136/bmjdrc-2023-003482.
45 Abraham TM, Pencina KM, Pencina MJ, Fox CS. Trends in Diabetes Incidence: The Framingham Heart

Study. *Diabetes Care.* 2015;38(3):482–487. doi:10.2337/dc14-1432
46 Siu AL. Screening for Abnormal Blood Glucose and Type 2 Diabetes Mellitus: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med*. 2015;163(11):861–868. doi:10.7326/M15-2345
47 Farmer R, Mathur R, Bhaskaran K, Eastwood SV, Chaturvedi N, Smeeth L. Promises and pitfalls of electronic health record analysis. *Diabetologia*. 2018;61(6):1241–1248. doi:10.1007/s00125-017-4518-6

SUPPORTING INFORMATION

Supplementary Table 1 Multivariable Cox regression coefficients (standard errors) of the type 2 diabetes mellitus risk prediction model by backward elimination method for predictors selection, Aichi, Japan, 2007–2019

Predictors	β (standard error)	*p* values
Body mass index, kg/m^2		
<23	Reference	
23–<25	0.499 (0.226)	0.027
25–<27.5	0.771 (0.245)	0.002
≧27.5	1.302 (0.275)	<0.001
Triglycerides, mg/dL		
<150	Reference	
≧150	0.313 (0.194)	0.107
High–density lipoprotein cholesterol, mg/dL		
<40	Reference	
≧40	–0.717 (0.267)	0.007
Fasting blood glucose, mg/dL		
<100	Reference	
100–<110	1.306 (0.191)	<0.001
110–<126	2.140 (0.253)	<0.001
SBP change trend, mmHg/year	0.029 (0.018)	0.101
SBP RMSE, mmHg		
T1	Reference	
T2	0.278 (0.234)	0.234
T3	0.579 (0.223)	0.010

SBP: systolic blood pressure
RMSE: root-mean-square error

Supplementary Table 2A Reclassification based on the 10-year predicted probability of T2DM incidence for event group

Predicted probability in Model 1	Predicted probability in Model 2		
	<0.05	<0.1	≧0.1
<0.05	79	5	0
<0.1	1	14	6
≧0.1	0	0	20

Event group was defined as the participants who developed T2DM within or equal to 10 years. The predictors in Model 1 and Model 2 are shown in Table 2.
T2DM: type 2 diabetes mellitus

Supplementary Table 2B Reclassification based on the 10-year predicted probability of T2DM incidence for nonevent group

Predicted probability in Model 1	Predicted probability in Model 2		
	<0.05	<0.1	≧0.1
<0.05	2166	79	1
<0.1	88	267	42
≧0.1	0	43	196

Nonevent group was defined as the participants who did not develop T2DM within or equal to 10 years. The predictors in Model 1 and Model 2 are shown in Table 2.
T2DM: type 2 diabetes mellitus

Supplementary Table 3 Comparisons of the prediction models with or without SBP RMSE in a sample that excluded participants who reported antihypertensive medication use from 2003 to 2007

	Model 1	Model 2	*p*
c-index (standard error)	0.798 (0.026)	0.800 (0.026)	0.93
Greenwood-Nam-D'Agostino test	2.46 (0.48)	3.60 (0.31)	
NRI_event[†], %	9.411		
NRI_nonevent[‡], %	0.168		
NRI_overall[§], %	9.579		<0.001
IDI, %	1.300		0.020

SBP: systolic blood pressure
RMSE: root-mean-square error
NRI: net reclassification improvement
IDI: integrated discrimination improvement
[†]NRI_event indicates the NRI calculated for event group.
[‡]NRI_nonevent indicates the NRI calculated for nonevent group.
[§]NRI_overall is the sum of NRI_event and NRI_nonevent indicates the total NRI for model 2 compared with Model 1.
Model 1 and Model 2 predictors are shown in Table 2.

Supplementary Table 4 Multivariable Cox regression HR (95% CI) of the type 2 diabetes mellitus risk prediction model in an analysis that censored participants at 60 years of age, Aichi, Japan, 2007–2019

Risk factors	Model 2 HR (95% CI)
Age, year	1.06 (1.02–1.10)
Sex	
Female	1 (reference)
Male	1.27 (0.67–2.42)

Body mass index, kg/m²	
<23.0	1 (reference)
23.0–<25.0	1.66 (0.99–2.77)
25.0–<27.5	2.98 (1.75–5.07)
≧27.5	3.86 (2.10–7.08)
Smoking status	
Non–current	1 (reference)
Current	1.22 (0.78–1.90)
Alcohol consumption, g/day	
0	1 (reference)
<23	0.72 (0.45–1.13)
23–<46	0.69 (0.34–1.41)
≧46	1.68 (0.71–3.99)
Regular exercise	
Yes	1 (reference)
No	1.05 (0.65–1.71)
Family history of diabetes	
No	1 (reference)
Yes	1.38 (0.50–3.81)
History of dyslipidemia	
No	1 (reference)
Yes	0.69 (0.30–1.55)
Baseline SBP, mmHg	1.00 (0.99–1.02)
Fasting blood glucose, mg/dL	
<100	1 (reference)
100–<110	3.40 (2.20–5.27)
110–<126	12.15 (6.98–21.14)
Triglycerides, mg/dL	
<150	1 (reference)
≧150	1.04 (0.67–1.59)
High–density lipoprotein cholesterol, mg/dL	
≧40	1 (reference)
<40	2.21 (1.22–4.03)
SBP change trend, mmHg/ year	1.02 (0.74–2.07)
SBP RMSE, mmHg	
T1	1 (reference)
T2	1.24 (0.74–2.07)
T3	1.68 (1.01–2.77)

HR: hazard ratio
CI: confidence interval
SBP: systolic blood pressure
RMSE: root-mean-square error

Supplementary Table 5 Comparisons of the prediction models with or without SBP RMSE in female participants

	Model 1	Model 2	*p*
c-index (standard error)	0.842 (0.056)	0.873 (0.042)	0.651
NRI_event[†], %	17.848		
NRI_nonevent[‡], %	–2.177		
NRI_overall[§], %	15.670		0.308
IDI, %	4.400		0.116

SBP: systolic blood pressure
RMSE: root-mean-square error
NRI: net reclassification improvement
IDI: integrated discrimination improvement
[†]NRI_event indicates the NRI calculated for event group.
[‡]NRI_nonevent indicates the NRI calculated for nonevent group.
[§]NRI_overall is the sum of NRI_event and NRI_nonevent indicates the total NRI for model 2 compared with Model 1.
Model 1 and Model 2 predictors are shown in Table 2.

Supplementary Table 6 Comparisons of the prediction models with or without SBP RMSE in male participants

	Model 1	Model 2	*p*
c-index (standard error)	0.771 (0.022)	0.775 (0.023)	0.923
NRI_event[†], %	6.134		
NRI_nonevent[‡], %	0.377		
NRI_overall[§], %	6.511		0.056
IDI, %	0.500		0.022

SBP: systolic blood pressure
RMSE: root-mean-square error
NRI: net reclassification improvement
IDI: integrated discrimination improvement
[†]NRI_event indicates the NRI calculated for event group.
[‡]NRI_nonevent indicates the NRI calculated for nonevent group.
[§]NRI_overall is the sum of NRI_event and NRI_nonevent indicates the total NRI for model 2 compared with Model 1.
Model 1 and Model 2 predictors are shown in Table 2.