

Development of a new scoring system for predicting the 5 year incidence of type 2 diabetes in Japan: the Toranomon Hospital Health Management Center Study 6 (TOPICS 6)

Y. Heianza · Y. Arase · S. D. Hsieh · K. Saito · H. Tsuji ·
S. Kodama · S. Tanaka · Y. Ohashi · H. Shimano ·
N. Yamada · S. Hara · H. Sone

Received: 28 February 2012 / Accepted: 9 August 2012 / Published online: 7 September 2012
© Springer-Verlag 2012

Abstract

Aims/hypothesis The aims of this study were to assess the clinical significance of introducing HbA_{1c} into a risk score for diabetes and to develop a scoring system to predict the 5 year incidence of diabetes in Japanese individuals.

Methods The study included 7,654 non-diabetic individuals aged 40–75 years. Incident diabetes was defined as fasting plasma glucose (FPG) ≥ 7.0 mmol/l, HbA_{1c} $\geq 6.5\%$ (48 mmol/mol) or self-reported clinician-diagnosed diabetes. We constructed a risk score using non-laboratory assessments (NLA) and evaluated improvements in risk prediction by adding elevated FPG, elevated HbA_{1c} or both to NLA.

Results The discriminative ability of the NLA score (age, sex, family history of diabetes, current smoking and BMI)

was 0.708. The difference in discrimination between the NLA + FPG and NLA + HbA_{1c} scores was non-significant (0.836 vs 0.837; $p=0.898$). A risk score including family history of diabetes, smoking, obesity and both FPG and HbA_{1c} had the highest discrimination (0.887, 95% CI 0.871, 0.903). At an optimal cut-off point, sensitivity and specificity were high at 83.7% and 79.0%, respectively. After initial screening using NLA scores, subsequent information on either FPG or HbA_{1c} resulted in a net reclassification improvement of 42.7% or 52.3%, respectively ($p<0.0001$). When both were available, net reclassification improvement and integrated discrimination improvement were further improved at 56.7% (95% CI 47.3%, 66.1%) and 10.9% (9.7%, 12.1%), respectively.

Electronic supplementary material The online version of this article (doi:10.1007/s00125-012-2712-0) contains peer-reviewed but unedited supplementary material, which is available to authorised users.

Y. Heianza · S. Kodama · H. Sone
Department of Internal Medicine,
Niigata University Faculty of Medicine,
Niigata, Japan

Y. Heianza · Y. Arase · S. D. Hsieh · K. Saito · H. Tsuji ·
S. Kodama · S. Hara · H. Sone (✉)
Health Management Center, Toranomon Hospital,
2-2-2 Toranomon,
Minato-ku, Tokyo 105-8470, Japan
e-mail: sone@med.niigata-u.ac.jp

Y. Heianza · H. Shimano · N. Yamada
Department of Internal Medicine,
University of Tsukuba Institute of Clinical Medicine,
Ibaraki, Japan

Y. Arase · S. D. Hsieh · H. Tsuji · S. Hara
Okinaka Memorial Institute for Medical Research,
Tokyo, Japan

K. Saito
Center for Medical Sciences,
Ibaraki Prefectural University of Health Sciences,
Ibaraki, Japan

S. Tanaka
Department of Clinical Trial Design & Management,
Translational Research Center, Kyoto University Hospital,
Kyoto, Japan

Y. Ohashi
Department of Biostatistics / Epidemiology
and Preventive Health Sciences, University of Tokyo,
Tokyo, Japan

Conclusions/interpretation Information on HbA_{1c} or FPG levels after initial screening by NLA can precisely refine diabetes risk reclassification.

Keywords HbA_{1c} · Prediction · Risk scores · Type 2 diabetes

Abbreviations

FPG	Fasting plasma glucose
IDI	Integrated discrimination improvement
NLA	Non-laboratory assessments
NRI	Net reclassification improvement
ROC	Receiver operating characteristic

Introduction

A recent systematic analysis showed that glycaemia and diabetes are rising global hazards, with the number of diabetic individuals having more than doubled over three decades [1]. Effective identification of individuals at high risk of developing diabetes is a major priority in managing this epidemic. Various rules have been developed to predict the incidence of diabetes in different ethnic groups [2–5]. Routine clinical markers available without laboratory testing have been shown to be predictive of the development of diabetes [6–18] and adding biochemical measures, in particular fasting plasma glucose (FPG), improves predictive accuracy [7, 8, 12–17]. On the other hand, adding complex data such as the results of oral glucose tolerance tests and measurements of insulin levels and inflammatory markers into a simple clinical model only minimally improves risk prediction while increasing cost and inconvenience [10, 19]. Similarly, adding genetic information to conventional risk factors does not appear to greatly refine the prediction of diabetes risk [14, 20, 21].

Introducing HbA_{1c} into a prediction model has been suggested to be effective in screening for future diabetes [14–16, 22–28], and a few models to predict the development of diabetes have recently been developed that concurrently include measurements of FPG, HbA_{1c} and other blood markers [15, 25, 29]. After revising the diagnostic criteria for type 2 diabetes by introducing HbA_{1c} in 2010 [30], the ADA published guidelines and recommendations for diagnosing diabetes based on published data or derived from expert consensus [31]. These guidelines emphasised again that ‘in addition to the long-standing criteria based on measurement of plasma glucose, diabetes can be diagnosed by demonstrating increased levels of HbA_{1c} concentrations’ [31]. Since the diagnostic criterion of HbA_{1c} ≥6.5% (48 mmol/mol) has been newly introduced not only by the ADA but also by the World Health Organization [32], the utility of HbA_{1c} testing as a screening tool to predict future diabetes, especially estimating risk among individuals below the diagnostic threshold of 6.5% (48 mmol/mol), should

be considered. In our previous study, we observed that elevations in the glycaemic markers FPG and HbA_{1c} were strongly predictive of the development of diabetes [33]. Therefore, it would be reasonable to consider whether their addition to other possible predictors would allow us to devise a simple, sensitive and specific algorithm to predict diabetes.

The new diagnostic criteria for diabetes based on HbA_{1c} do not require a fasting sample [30, 31]. This benefit deserves consideration in the development of a new scoring system by quantifying the additional predictive power of adding HbA_{1c} to a risk score based on non-laboratory assessments (NLA). A recent study of older British men and women described a two-stage scoring system that combined a simple clinical assessment as the first step, with subsequent inclusion of HbA_{1c} and other biochemical measures that could be assessed in a non-fasting state as the second step [16]. However, it has not been clarified to what extent available information on elevated levels of either HbA_{1c}, FPG or both would improve overall predictive accuracy and risk reclassification after initial screening by a simple NLA score to establish a two-stage model. Whether HbA_{1c} and FPG are equivalent tools for prediction in risk scores and whether their combination can provide a distinct advantage are important and practical questions at the present time, particularly given the energy expended and emphasis placed on establishing programmes for primary prevention. Therefore, in this study we aimed to assess the clinical significance of introducing HbA_{1c} into a risk score and to develop and validate a scoring system to predict the 5 year incidence of diabetes in Japanese men and women. We evaluated the performance of the various diabetes risk scores that we developed as two-stage scoring systems to predict the 5 year incidence of type 2 diabetes.

Methods

Study population The Toranomon Hospital Health Management Center Study (TOPICS) included a cohort consisting mainly of apparently healthy Japanese government employees who underwent annual examinations for health screening, in addition to some participants from the general public. All persons were interviewed at each examination using standard questionnaires that gathered information on demographic characteristics, health-related habits and medical history. A total of 29,584 individuals underwent a baseline health examination during the period from 1997 to 2002. To avoid overlapping of data from the same individuals who had health examinations in multiple years, we used data only from the first visit. Enrolled in this study were 9,344 individuals aged 24–82 years who underwent a re-examination 5 years after the initial examination. Individuals who had diabetes at the baseline examination ($n=397$) or with missing data on baseline characteristics ($n=137$) were

then excluded. For reasons of clinical significance, we curtailed the current analysis to individuals aged 40–75 years, an age group that is the most likely to be screened for risk of diabetes. Consequently 1,148 individuals (12% of the 9,344 individuals) under the age of 40 years and 28 over the age of 75 years were excluded. Twenty individuals who had both missing data and baseline diabetes or were excluded because of age were therefore counted twice; thus, based on the above-described study design, overall 18% of individuals ($n=1,690$) were excluded from the 9,344 individuals originally considered as study participants.

After these adjustments, 7,654 individuals (2,211 women and 5,443 men) aged 40–75 years comprised the cohort for derivation of the risk score to predict the 5 year incidence of diabetes. Diagnosis of type 2 diabetes was made according to the ADA criteria of an FPG level of 7.0 mmol/l or higher, self-reported clinician-diagnosed diabetes or HbA_{1c} of 6.5% (48 mmol/mol) or higher [30]. Informed consent was obtained from all participants. The study protocol followed the Japanese government's Ethical Guidelines Regarding Epidemiological Studies in accordance with the Declaration of Helsinki and was reviewed by the institutional review board at Toranomon Hospital.

Assessment of risk factors Height and weight were measured without shoes or heavy clothing, and BMI (kg/m^2) was calculated. Blood pressure was measured by trained hospital staff with the participant in a seated position. Individuals with a systolic blood pressure of 140 mmHg or higher or a diastolic blood pressure of 90 mmHg or higher, or who were under medical treatment, were considered to have hypertension. Current smoking habit, first-degree relatives (i.e., parent or sibling) with diabetes and self-reported histories of dyslipidaemia and cardiovascular disease were assessed using a standard questionnaire. Blood samples were collected after an overnight fast (12 h) and measured using an automatic clinical chemistry analyser (LABO-SPECT 008, Hitachi, Tokyo, Japan). Blood glucose was measured by enzymatic methods and HbA_{1c} was assessed by high-performance liquid chromatography. The value for HbA_{1c} was estimated as the National Glycohemoglobin Standardization Program value (%), calculated by the formula $\text{HbA}_{1c} (\%) = 1.02 \times \text{HbA}_{1c} (\text{Japan Diabetes Society}) (\%) + 0.25\%$ [34].

Validation of the developed risk scores The validity of the derived risk score was tested in a separate study population that underwent a first examination in 2003, 2004 or 2005, with prospective follow-up for 5 years ($n=1,976$). Among the 1,976 individuals, we excluded those with diabetes ($n=106$), younger than 40 years ($n=424$) or older than 75 years ($n=2$), or with missing data ($n=15$); thus, 1,437 individuals

aged 40–75 years were included in the validation study. During the 5 year follow-up, we documented 57 incident cases of diabetes for an incidence of 4.0%.

Statistical analysis A logistic regression model was used to investigate independent predictors of the incidence of diabetes, and β coefficients, ORs and their 95% CIs were estimated. We initially tested all the variables in the univariate regression model to determine which were significantly predictive of the development of diabetes. In model building, we analysed data using a multiple regression model with both backward and forward elimination methods from the initial model until we reached a final model to select independent and significant predictors without including FPG and HbA_{1c} values. After the simplicity of the screening algorithms was also considered, we developed the simplest NLA model that was considered to have reasonable predictive abilities by performing logistic regression analysis with the forward selection method and a significance level of 0.05. The principal criteria for selecting variables in the developed NLA model included results of the Hosmer–Lemeshow goodness-of-fit test and changes in the area under the receiver operating characteristic (ROC) curve. In addition, whether information in the NLA model was widely available and would be known to participants without consulting a medical professional was also considered in the model building.

To evaluate whether predictive ability was improved by adding FPG, HbA_{1c} or both into the NLA model, we compared discriminative ability by calculating the area under the ROC curve, and statistical significance was tested using the method of DeLong and colleagues [35]. We also calculated net reclassification improvement (NRI) using three risk categories (<5%, 5–15% and >15%) and integrated discrimination improvement (IDI) [36]. To develop the risk scores for predicting the 5 year incidence of diabetes, we estimated point scores from the β coefficients of the multivariate logistic regression analysis. Interactions between sex and key variables in the score (obesity, current smoking habit, impaired fasting glucose or elevated HbA_{1c} values) were also tested but significant interactions ($p<0.05$) were not identified. Analysis was performed with IBM SPSS Statistics version 19 (IBM, Armonk, NY, USA) and STATA software version 11 (STATA Corp., College Station, TX, USA). Statistical significance was considered for $p<0.05$.

Results

Derivation of diabetes risk score During a 5 year follow-up period, we documented 289 incident cases of diabetes. Non-laboratory measurements of age, male sex, family history of

diabetes, current smoking habit, BMI, resting heart rate, hypertension and self-reported history of dyslipidaemia were significantly associated with the development of diabetes in the univariate logistic regression model (Table 1). In the multiple regression model with the forward elimination method, BMI was most strongly predictive of future diabetes (Hosmer–Lemeshow test, $p=0.428$; area under the ROC 0.653 [95% CI 0.620, 0.687]). Results of calibration by the Hosmer–Lemeshow test and discrimination of the model were $p=0.304$ and area under the ROC curve 0.722 (95% CI 0.694, 0.750), respectively, after entering the four variables of family history of diabetes, sex, age and current smoking habit (NLA model in Table 2). In the final step, further adding data on resting heart rate and hypertension minimally improved calibration (Hosmer–Lemeshow test, $p=0.19$), discrimination (0.732; 95% CI 0.705, 0.759), NRI (2.4%; 95% CI -1.3, 6.1) and IDI (0.2%; 95% CI 0.05, 0.4). Considering these results and simplicity for use in routine care settings, an NLA model was constructed with the five non-laboratory markers that could be assessed before clinical measurements were performed.

Predicting the development of diabetes by adding either FPG (model 2) or HbA_{1c} (model 3) into the NLA model significantly ($p<0.0001$) improved the area under the ROC curve to a similar degree (Table 2). No significant difference was observed in the discriminative ability between model 2 and model 3 ($p=0.435$). When the addition of data on HbA_{1c}, HDL-cholesterol and alanine aminotransferase that

could be assessed in a non-fasting state into the NLA model (detailed data not shown) was analysed, the area under the ROC curve (0.864; 95% CI 0.842, 0.887), NRI (47.2%; 95% CI 38.0, 56.3%) and IDI (14.7%; 95% CI 12.4, 17.0%) indicated refined risk prediction, but similar to that of NLA + HbA_{1c} (model 3) and NLA + FPG (model 2). When we added both FPG and HbA_{1c} into the NLA model (model 4), the area under the ROC curve was improved ($p<0.0001$) to 0.907 (95% CI 0.890, 0.925), which was the highest of the four models ($p<0.0001$). NRI and IDI were also improved by adding information on FPG and HbA_{1c} either singly or in combination into model 1 ($p<0.0001$).

Then we developed four diabetes risk scores using the significant predictors in each model by categorising the variables for practical use in clinical settings (Table 3). Calibration by the Hosmer–Lemeshow test was $p\geq 0.05$ for all of the models and showed a reasonable fit (see electronic supplementary material [ESM] Fig. 1). The estimated probability of developing diabetes 5 years later gradually escalated in association with higher risk scores (ESM Table 1). The area under the ROC curve of the NLA score was 0.708 (95% CI 0.679, 0.737); this risk score with omission of glycaemic measurements discriminated relatively well between individuals who did and did not develop diabetes (ESM Fig. 2). The two scores that included either elevated FPG levels or elevated HbA_{1c} levels had greatly improved discrimination. The area under the ROC curve for the NLA + FPG + HbA_{1c} score generated

Table 1 Baseline characteristics of the participants and ORs for each variable to predict the 5 year incidence of diabetes

Variable	Derivation cohort		Validation cohort
	Total participants ($n=7,654$)	Univariate model, OR (95% CI)	Total participants ($n=1,437$)
Age (years)	50.2±6.9	1.02 (1.001, 1.03)	49.2±7.6
Male sex	5,443 (71.1)	2.39 (1.72, 3.31)	984 (68.5)
Family history of diabetes (yes) ^a	1,281 (16.7)	2.68 (2.09, 3.45)	265 (18.4)
Current smoking habit (yes)	2,019 (26.4)	1.57 (1.23, 2.01)	338 (23.5)
BMI (kg/m ²)	22.8±2.8	1.21 (1.17, 1.26)	23.0±3.0
Resting heart rate (beats/min)	65±9	1.02 (1.01, 1.03)	65 (9)
Hypertension ^b	1,582 (20.7)	1.97 (1.53, 2.53)	284 (19.8)
Self-reported history of dyslipidaemia	386 (5.0)	1.84 (1.20, 2.81)	67 (4.7)
HDL-cholesterol (mmol/l)	1.41±0.40	0.33 (0.24, 0.47)	1.49±0.34
Alanine aminotransferase (U/l)	25±17	1.02 (1.02, 1.03)	25±26
Self-reported history of coronary heart disease	83 (1.1)	1.64 (0.66, 4.09)	17 (1.2)
FPG (mmol/l)	5.3±0.5	19.8 (15.3, 25.7)	5.3±0.5
HbA _{1c} (%)	5.2±0.3	7.56 (6.24, 9.14) ^c	5.3±0.3
HbA _{1c} (mmol/mol)	34±4	1.45 (1.40, 1.50)	34±4

Data are n (%) or mean ± SD unless indicated otherwise

^aHistory of parents or siblings having diabetes

^bHypertension defined as systolic blood pressure ≥ 140 mmHg, diastolic blood pressure ≥ 90 mmHg or medical treatment

^cHbA_{1c} 0.5% increment

Table 2 Prediction models for 5 year incidence of type 2 diabetes using NLA, FPG and HbA_{1c}

Variable	Model 1: NLA		Model 2: NLA + FPG		Model 3: NLA + HbA _{1c}		Model 4: NLA + FPG and HbA _{1c}	
	OR (95% CI)	<i>p</i> value	OR (95% CI)	<i>p</i> value	OR (95% CI)	<i>p</i> value	OR (95% CI)	<i>p</i> value
Age (years)	1.03 (1.02, 1.05)	<0.001	1.02 (1.01, 1.04)	0.011	0.99 (0.97, 1.01)	0.312	0.99 (0.97, 1.02)	0.593
Male sex (yes)	1.88 (1.34, 2.64)	<0.001	0.90 (0.63, 1.31)	0.594	1.84 (1.28, 2.63)	<0.001	1.02 (0.70, 1.50)	0.905
Family history of diabetes (yes) ^a	2.83 (2.19, 3.65)	<0.001	2.00 (1.50, 2.65)	<0.001	2.07 (1.56, 2.73)	<0.001	1.66 (1.23, 2.25)	<0.001
Current smoking (yes)	1.50 (1.16, 1.93)	0.002	1.67 (1.26, 2.21)	<0.001	1.26 (0.95, 1.67)	0.102	1.53 (1.14, 2.06)	0.005
BMI (kg/m ²)	1.21 (1.16, 1.26)	<0.001	1.10 (1.05, 1.15)	<0.001	1.15 (1.10, 1.20)	<0.001	1.08 (1.03, 1.13)	0.002
FPG (mmol/l)	–	–	17.1 (13.0, 22.5)	<0.001	–	–	9.05 (6.74, 12.2)	<0.001
HbA _{1c} (%) ^b	–	–	–	–	6.97 (5.71, 8.51)	<0.001	4.06 (3.27, 5.04)	<0.001
Calibration ^c	0.304		0.188		0.140		0.138	
AUCROC (95% CI)	0.722 (0.694, 0.750)		0.870 (0.848, 0.893)		0.858 (0.836, 0.881)		0.907 (0.890, 0.925)	
Change in AUCROC, <i>p</i> value	Ref.		<0.0001		<0.0001		<0.0001	
NRI (%) (95% CI)	Ref.		53.1 (43.5, 62.7) [†]		47.4 (38.2, 56.5) [†]		58.3 (48.6, 68.0) [†]	
IDI (%) (95% CI)	Ref.		14.6 (12.5, 16.8) [†]		13.2 (11.1, 15.2) [†]		21.7 (18.9, 24.6) [†]	

^a History of parents or siblings having diabetes

^b Hosmer–Lemeshow test, *p* values

^c HbA_{1c} 0.5% increment

[†] *p*<0.0001

AUCROC, area under the ROC curve; Ref., reference

Table 3 Development of diabetes risk score to predict 5 year incidence of type 2 diabetes using NLA, FPG and HbA_{1c}

Variable	Incident cases/ <i>n</i>	NLA score			NLA + FPG score			NLA + HbA _{1c} score			NLA + FPG + HbA _{1c} score		
		β	OR (95% CI)	P	β	OR (95% CI)	P	β	OR (95% CI)	P	β	OR (95% CI)	P
Age ≥ 55 years	72/1,715	0.334	1.40 (1.06, 1.85)	2	0.325	1.38 (1.04, 1.85)	1	–	–	–	–	–	–
Male sex	246/5,443	0.646	1.91 (1.35, 2.69)	4	–	–	–	0.742	2.10 (1.49, 2.97)	3	–	–	–
Family history of diabetes (yes) ^a	98/1,281	1.022	2.78 (2.15, 3.59)	7	0.791	2.21 (1.69, 2.87)	3	0.819	2.27 (1.74, 2.96)	3	0.661	1.94 (1.47, 2.55)	2
Current smoking habit (yes)	103/2,019	0.381	1.46 (1.13, 1.89)	2	0.417	1.52 (1.17, 1.97)	2	–	–	–	0.353	1.42 (1.09, 1.86)	1
BMI (kg/m ²)													
<23	76/3,520	–	1.00 (Ref.)	0	–	1.00 (Ref.)	0	–	1.00 (Ref.)	0	–	1.00 (Ref.)	0
23–24	76/2,202	0.366	1.44 (1.04, 2.00)	2	0.138	1.15 (0.82, 1.60)	1	0.268	1.31 (0.93, 1.83)	1	0.083	1.09 (0.77, 1.53)	0
≥ 25	137/1,932	1.103	3.01 (2.24, 4.04)	7	0.678	1.97 (1.46, 2.66)	3	0.856	2.35 (1.74, 3.19)	3	0.544	1.72 (1.26, 2.35)	2
FPG (mmol/l)													
<5.3	26/3,894	–	–	–	–	–	–	–	–	–	–	–	–
5.3–5.5	30/1,655	–	–	–	0.931	2.54 (1.49, 4.31)	4	–	–	–	0.680	1.97 (1.15, 3.38)	2
5.6–6.9	233/2,105	–	–	–	2.711	15.0 (9.93, 22.8)	11	–	–	–	2.244	9.43 (6.18, 14.4)	7
HbA _{1c}													
<5.2% (<33 mmol/mol)	21/3,500	–	–	–	–	–	–	–	–	–	–	–	–
5.2–5.5% (33–37 mmol/mol)	66/2,556	–	–	–	–	–	–	1.430	4.18 (2.55, 6.85)	5	1.174	3.23 (1.96, 5.33)	4
$\geq 5.6\%$ (38–46 mmol/mol)	202/1,598	–	–	–	–	–	–	3.034	20.8 (13.2, 32.8)	11	2.541	12.7 (8.00, 20.2)	8
Intercept		–4.684			–5.555			–6.192			–6.569		
Calibration ^b		0.058			0.743			0.979			0.174		

^a History of parents or siblings having diabetes^b Hosmer–Lemeshow test, *p* values

P, points scored; Ref., reference

the highest discrimination of 0.887 (95% CI 0.871, 0.903) among the four risk scores ($p < 0.0001$). When we investigated screening performance according to cut-off values (Table 4), the NLA score of 6 points or higher had a sensitivity of 89.6% but specificity of only 32.3%. The ‘NLA + FPG score’ with 7 points or higher was associated with a sensitivity of 89.6% and high specificity at 63.9% and the ‘NLA + HbA_{1c} score’ with 9 points or higher had a sensitivity of 86.9% and specificity of 63.1%. We found that the ‘NLA + FPG + HbA_{1c} score’ with 11 points or higher generated the highest combination of sensitivity (83.7%) and specificity (79.0%) and had the best positive predictive value (13.5%) and positive likelihood ratio (3.99) among the risk scores. We also confirmed

that the screening performance of the four diabetes risk scores was well validated in a population ($n = 1,437$) separate from the derivation cohort.

Table 5 shows the improved discriminative ability and risk reclassification provided by the addition of FPG and/or HbA_{1c} after screening by the NLA risk score. Results in which we calculated the NRI showed that additional information on elevated levels of either FPG or HbA_{1c} appropriately reclassified participants into predicted 5 year risk categories. When both FPG and HbA_{1c} were available after the initial screening by the NLA risk score, 56% of individuals who developed diabetes were appropriately reclassified and the NRI was 56.7% (95% CI 47.3%, 66.1%). The IDI

Table 4 Screening performance of the developed diabetes risk scores for predicting future type 2 diabetes

Score	Area under the ROC curve (95% CI)	Cut-off value	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Likelihood ratio (+)	Youden index (%)
Derivation cohort ($n = 7,654$)								
NLA score	0.708 (0.679, 0.737)	≥6	89.6	32.3	4.9	98.8	1.32	21.9
		≥7	77.9	53.2	6.1	98.4	1.66	31.1
		≥8	74.4	57.6	6.4	98.3	1.76	32.0
NLA + FPG score	0.836 (0.815, 0.856)	≥7	89.6	63.9	8.9	99.4	2.48	53.5
		≥10	82.0	73.1	10.7	99.0	3.05	55.1
		≥11	81.3	74.2	11.0	99.0	3.16	55.6
NLA + HbA _{1c} score	0.837 (0.815, 0.859)	≥9	86.9	63.1	8.5	99.2	2.35	49.9
		≥11	82.0	71.3	10.1	99.0	2.86	53.3
		≥12	70.9	81.3	12.9	98.6	3.79	52.2
NLA + FPG + HbA _{1c} score	0.887 (0.871, 0.903)	≥9	90.3	70.6	10.8	99.5	3.08	60.9
		≥10	86.5	74.7	11.8	99.3	3.42	61.2
		≥11	83.7	79.0	13.5	99.2	3.99	62.7
Validation cohort ($n = 1,437$)								
NLA score	0.727 (0.670, 0.784)	≥6	94.7	32.7	5.5	99.3	1.41	27.4
		≥7	82.5	50.0	6.4	98.6	1.65	32.5
		≥8	78.9	55.4	6.8	98.5	1.77	34.4
NLA + FPG score	0.865 (0.831, 0.900)	≥7	96.5	59.9	9.0	99.8	2.41	56.4
		≥10	93.0	69.9	11.3	99.6	3.08	62.8
		≥11	93.0	71.1	11.7	99.6	3.22	64.1
NLA + HbA _{1c} score	0.879 (0.842, 0.916)	≥9	94.7	58.6	8.6	99.6	2.29	53.3
		≥11	91.2	67.5	10.4	99.5	2.81	58.8
		≥12	84.2	79.6	14.5	99.2	4.12	63.8
NLA + FPG + HbA _{1c} score	0.913 (0.878, 0.947)	≥9	94.7	66.3	10.4	99.7	2.81	61.0
		≥10	94.7	70.1	11.6	99.7	3.17	64.8
		≥11	94.7	75.5	13.8	99.7	3.87	70.2

NLA score: age ≥55 years (2 points) + male sex (4 points) + history of parents or siblings having diabetes (7 points) + current smoking habit (2 points) + BMI 23–24 kg/m² (2 points) or ≥25 kg/m² (7 points)

NLA + FPG score: age ≥55 years (1 point) + history of parents or siblings having diabetes (3 points) + current smoking habit (2 points) + BMI 23–24 kg/m² (1 point) or ≥25 kg/m² (3 points) + FPG 5.3–5.5 mmol/l (4 points) or 5.6–6.9 mmol/l (11 points)

NLA + HbA_{1c} score: male sex (3 points) + history of parents or siblings having diabetes (3 points) + BMI 23–24 kg/m² (1 point) or ≥25 kg/m² (3 points) + HbA_{1c} 5.2–5.5% (33–37 mmol/mol) (5 points) or 5.6–6.4% (38–46 mmol/mol) (11 points)

NLA + FPG + HbA_{1c} score: history of parents or siblings having diabetes (2 points) + current smoking habit (1 point) + BMI ≥25 kg/m² (2 points) + FPG 5.3–5.5 mmol/l (2 points) or 5.6–6.9 mmol/l (7 points) + HbA_{1c} 5.2–5.5% (33–37 mmol/mol) (4 points) or 5.6–6.4% (38–46 mmol/mol) (8 points)

Table 5 Reclassification of 5 year predicted risk and change in risk discrimination for future type 2 diabetes after addition of glycaemic markers to a risk score that includes NLA

Variable	NLA + FPG score	NLA + HbA _{1c} score	NLA + FPG + HbA _{1c} score
Discrimination			
Change in area under the ROC curve ^a	0.128	0.129	0.179
<i>p</i> value	<0.0001	<0.0001	<0.0001
<i>p</i> value ^b	Reference	0.898	<0.0001
Reclassification			
Participants who developed diabetes by 5 years (<i>n</i> =289)			
Appropriately reclassified, <i>n</i> (%)	154 (53)	176 (61)	163 (56)
Inappropriately reclassified, <i>n</i> (%)	27 (9)	37 (13)	27 (9)
No change, <i>n</i> (%)	108 (37)	76 (26)	99 (34)
Participants who had not developed diabetes by 5 years (<i>n</i> =7,365)			
Appropriately reclassified, <i>n</i> (%)	1,282 (17)	1,470 (20)	1,421 (19)
Inappropriately reclassified, <i>n</i> (%)	1,376 (19)	1,161 (16)	711 (10)
No change, <i>n</i> (%)	4,707 (64)	4,734 (64)	5,233 (71)
NRI % (95% CI)	42.7 (33.4, 51.9)	52.3 (42.3, 62.3)	56.7 (47.3, 66.1)
<i>p</i> value	<0.0001	<0.0001	<0.0001
IDI, % (95% CI)	4.9 (4.3, 5.5)	6.6 (5.8, 7.4)	10.9 (9.7, 12.1)
<i>p</i> value	<0.0001	<0.0001	<0.0001

^aReference was area under the ROC curve of 0.708 (95% CI 0.679, 0.737) for the NLA score

^b*p* value for changes compared with the NLA + FPG score

was also the highest, that is, 10.9% (95% CI 9.7%, 12.1%), when FPG and HbA_{1c} were introduced into the NLA risk score at the second screening test.

Discussion

Using information on family history of diabetes, obesity, current smoking habit and elevated levels of FPG and HbA_{1c} within a non-diabetic range, we developed a simple, highly sensitive and specific scoring system to predict the 5 year risk of developing diabetes. Introducing HbA_{1c} into a risk score that included FPG and non-laboratory measurements further refined the risk prediction and contributed to creating a valid and simple scoring system. In the past, various assessment tools have been developed for diverse ethnic groups [2–5] and the screening accuracy of some of these instruments has also been validated in external populations [6–11, 13, 18, 19, 37–39]. However, none of those validated scoring systems concurrently used HbA_{1c} and FPG with other routinely available clinical markers to derive a feasible scoring system to predict risk of diabetes. Our study provides four diabetes risk algorithms that can be used both within and outside of clinical practice and in non-fasting and fasting states, and found that in this two-stage scoring system, after initial screening by the developed NLA risk score, subsequent available information on FPG, HbA_{1c} or both precisely refined diabetes risk reclassification in Japanese men and women.

Combining information on FPG or impaired fasting glucose with a simple diabetes risk score has been reported to increase predictive ability [7, 8, 12–17]. A study reported that screening models using the combination of HbA_{1c}, BMI and FPG accurately identified individuals at risk of future clinically diagnosed diabetes [22], although the factors that remained significant were different from those found in the present study. The EPIC-Potsdam Study reported that discrimination by the German Diabetes Risk Score (including anthropometry and lifestyle characteristics) [9] had an area under the ROC curve of 0.8465, which was improved by adding FPG (0.8672), HbA_{1c} (0.8859) or both (0.8926) [14]. Our simple risk scoring systems with non-laboratory measures that include elevated FPG or HbA_{1c} levels have similarly excellent discrimination with no significant difference. Results of the EPIC-Potsdam Study also suggested that risk reclassification was improved by adding FPG (IDI 0.0553) or HbA_{1c} (IDI 0.0974) values using the prediction model by the German Diabetes Risk Score as a reference [14]. Recent prospective studies have examined the utility of introducing HbA_{1c} testing for predicting diabetes [15, 16, 23–27], and some reports have described the development of models to predict future diabetes using blood variables including FPG and HbA_{1c} [15, 25, 29]. More recently and concomitantly with the preparation of this manuscript, a risk score that concurrently included measurements of FPG, HbA_{1c} and other biochemical markers was reported in a Korean population [40]. However, whether HbA_{1c} and

FPG are equivalent tools for prediction in risk scores, and whether their combination provides a distinct advantage, are important and practical questions at the present time, particularly given the energy expended and emphasis placed on establishing programmes for primary prevention. Our study suggests that adding elevated HbA_{1c} levels within a non-diabetic range into a diabetes risk score that includes non-laboratory measurements improves its screening performance, and that HbA_{1c} and FPG equally contribute in a major way to our risk scoring system.

Wannamethee et al developed a risk score using simple blood markers and HbA_{1c} values that were not dependent on a fasting state and could therefore be used at any time of day [16]. They also suggested that an approach using simple clinical assessments in the first instance followed by the use of routine blood markers that do not require a fasting sample might be cost-effective in identifying those most likely to be at high risk for diabetes [16]. However, it has not been clarified whether adding both FPG and HbA_{1c} into risk scoring systems will generate a distinct advantage for improving overall screening performance. We found that after initial screening using a screening score based on non-laboratory measurements, additional information on either elevated FPG levels or elevated HbA_{1c} levels within a non-diabetic range improved risk prediction to a similar extent in our study population. Although which score will be most useful for predicting diabetes risk will depend on the individual's situation and the particular screening setting, our study results suggest that if both glycaemic data are available to add to the NLA risk score, clinicians might more precisely estimate the risk of diabetes development and take measures with their patients to prevent onset of the disease.

Although whether the developed tools might be useful beyond the Japanese population must be evaluated in other ethnic groups, our scores included common risk factors for diabetes, such as age, family history of diabetes, sex and obesity, compared with components of previously developed risk scores in other ethnicities [2–4]. A current smoking habit as a modifiable factor was also included in our developed scores. A meta-analysis has shown that being an active smoker is associated with a 1.44-fold higher risk of developing diabetes compared with being a non-smoker [41]. According to a recent study that reported the DETECT-2 update of the Finnish diabetes risk questionnaire, adding information on smoking and family history of diabetes into the original Finnish risk questionnaire improved its predictive ability [42]. Our results indicate that control of these two modifiable risk factors, that is, obesity and current smoking, should be given priority in preventing diabetes regardless of glycaemic status. Nonetheless, as to the magnitude of the effect of each risk factor on the incidence of diabetes, on average the BMI in Asian individuals is lower than that in other populations [43], and Asian

individuals tend to develop diabetes at lower BMI levels than white individuals [44]. It also should be considered that results of performance of HbA_{1c} as a screening test might differ according to race [45]; therefore, whether the thresholds shown in this study are universally useful should be validated in other ethnic groups.

Since the utility of HbA_{1c} as a screening tool for individuals at risk of diabetes has been recognised [46, 47], it is reasonable that the use of two glycaemic indicators would be considered in developing an effective screening score for risk of diabetes. Our findings suggest that a scoring system should not be limited to using data from key conventional variables, as reported in the past, but should also include HbA_{1c} in the development of a diabetes risk score to further improve screening for future diabetes. However, measuring both increases the cost of screening, and the cost-effectiveness of the use of HbA_{1c} in a two-stage test or as a simultaneous test with FPG in screening to predict diabetes should be considered. Nonetheless, studies confirm that conducting screening is more cost-effective than not performing screening, both from the health system and societal perspectives [48, 49]. An analysis of cost-effectiveness showed that in a comparison with no screening with simulated screening strategies, screening would theoretically reduce the incidence of microvascular complications and myocardial infarction and increase the number of quality-adjusted life-years over a follow-up period of 50 years [49]. A recent validation study of previously developed diabetes risk scores indicated that a risk score by Kahn et al that included laboratory measurements [13] had a high discriminatory value [18]. It was also reported that the Finnish diabetes risk score without blood testing [6] might be a more practical and less expensive screening test [18], although additional costs resulting from the necessary follow-up of individuals with a false-positive test result would be a consideration [50]. Performance of our risk scores with a true two-stage step-wise strategy would need to be evaluated in future research.

One strength of our study is the availability of data on a large number of individuals over a lengthy period. An additional strength is that this study addresses an important topic that will contribute to more accurate risk classification and allow strategies to be implemented to prevent the development of diabetes. Several limitations must be considered. Since our cohort consisted of Japanese individuals who had annual health check-ups, the generalisability of our results should be investigated in general populations. Our cohort included mainly men (71% of total participants) and the database did not allow us to conduct sex-stratified analyses to develop separate risk scores for men and women. In addition, from this prospective study we cannot suggest that our developed NLA score might be used as a self-assessment tool to identify individuals with presently unknown diabetes.

In conclusion, our study demonstrates that by adding information on HbA_{1c} into a risk scoring system that included FPG, family history of diabetes, current smoking habit and obesity, we could develop a simple, sensitive and specific algorithm that was clinically relevant to predict the 5 year risk of diabetes. Our risk score may contribute to predicting the future development of diabetes and thereby identifying individuals who will likely benefit from early interventions.

Acknowledgements All authors sincerely thank the late Professor and Director K. Kosaka, who established the foundation for the TOPICS project and represented the foremost pillar of spiritual support for the TOPICS project. Thanks are also extended to S. Fukuya at the University of Tsukuba Institute of Clinical Medicine for her excellent secretarial assistance. Parts of this study were presented at the 72nd Scientific Sessions of the ADA, 11 June 2012, Philadelphia, PA.

Funding This work was supported in part by the Ministry of Health, Labour and Welfare, Japan. Y. Heianza and H. Sone are recipients of a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science. The sponsor had no role in the design or conduct of the study.

Duality of interest The authors declare that there is no duality of interest associated with this manuscript.

Contribution statement YH contributed to the study concept and design, acquisition of data, analysis and interpretation of data, drafting and revising the manuscript, statistical analysis and study supervision. YA, HT, SDH and SH contributed to acquisition of data, technical or material support, and drafting and revising the manuscript. SK and KS contributed to acquisition of data, analysis and interpretation of data, statistical analysis and drafting and revising the manuscript. ST and YO contributed to statistical analysis, interpretation of data and revising the manuscript. HSI and NY contributed to the acquisition of data and material support and drafting and revising the manuscript. HSo contributed to the study concept and design, acquisition of data, analysis and interpretation of data, drafting and revising the manuscript, statistical analysis, funding and study supervision. All authors were involved in the writing of the manuscript and approved the final version of this article.

References

- Danaei G, Finucane MM, Lu Y et al (2011) National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. *Lancet* 378:31–40
- Buijsse B, Simmons RK, Griffin SJ, Schulze MB (2011) Risk assessment tools for identifying individuals at risk of developing type 2 diabetes. *Epidemiol Rev* 33:46–62
- Noble D, Mathur R, Dent T, Meads C, Greenhalgh T (2011) Risk models and scores for type 2 diabetes: systematic review. *BMJ* 343:d7163
- Collins GS, Mallett S, Omar O, Yu LM (2011) Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. *BMC Med* 9:103
- Sasai H, Sairenchi T, Irie F, Iso H, Tanaka K, Ota H (2008) Development of a diabetes risk prediction sheet for specific health guidance. *Nippon Koshu Eisei Zasshi* 55:287–294 [article in Japanese]
- Lindstrom J, Tuomilehto J (2003) The diabetes risk score: a practical tool to predict type 2 diabetes risk. *Diabetes Care* 26:725–731
- Schmidt MI, Duncan BB, Bang H et al (2005) Identifying individuals at high risk for diabetes: the Atherosclerosis Risk in Communities study. *Diabetes Care* 28:2013–2018
- Aekplakorn W, Bunnag P, Woodward M et al (2006) A risk score for predicting incident diabetes in the Thai population. *Diabetes Care* 29:1872–1877
- Schulze MB, Hoffmann K, Boeing H et al (2007) An accurate risk score based on anthropometric, dietary, and lifestyle factors to predict the development of type 2 diabetes. *Diabetes Care* 30:510–515
- Wilson PW, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB Sr (2007) Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. *Arch Intern Med* 167:1068–1074
- Balkau B, Lange C, Fezeu L et al (2008) Predicting diabetes: clinical, biological, and genetic approaches: data from the Epidemiological Study on the Insulin Resistance Syndrome (DESIR). *Diabetes Care* 31:2056–2061
- Gao WG, Qiao Q, Pitkaniemi J et al (2009) Risk prediction models for the development of diabetes in Mauritian Indians. *Diabet Med* 26:996–1002
- Kahn HS, Cheng YJ, Thompson TJ, Imperatore G, Gregg EW (2009) Two risk-scoring systems for predicting incident diabetes mellitus in U.S. adults age 45 to 64 years. *Ann Intern Med* 150:741–751
- Schulze MB, Weikert C, Pischon T et al (2009) Use of multiple metabolic and genetic markers to improve the prediction of type 2 diabetes: the EPIC-Potsdam Study. *Diabetes Care* 32:2116–2119
- Rathmann W, Kowall B, Heier M et al (2010) Prediction models for incident type 2 diabetes mellitus in the older population: KORA S4/F4 cohort study. *Diabet Med* 27:1116–1123
- Wannamethee SG, Papacosta O, Whincup PH et al (2011) The potential for a two-stage diabetes risk algorithm combining non-laboratory-based scores with subsequent routine non-fasting blood tests: results from prospective studies in older men and women. *Diabet Med* 28:23–30
- Doi Y, Ninomiya T, Hata J et al (2012) Two risk score models for predicting incident type 2 diabetes in Japan. *Diabet Med* 29:107–114
- Schmid R, Vollenweider P, Bastardot F, Waeber G, Marques-Vidal P (2012) Validation of 7 type 2 diabetes mellitus risk scores in a population-based cohort: CoLaus study. *Arch Intern Med* 172:188–189
- Stern MP, Williams K, Haffner SM (2002) Identification of persons at high risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? *Ann Intern Med* 136:575–581
- Meigs JB, Shrader P, Sullivan LM et al (2008) Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med* 359:2208–2219
- Talmud PJ, Hingorani AD, Cooper JA et al (2010) Utility of genetic and non-genetic risk factors in prediction of type 2 diabetes: Whitehall II prospective cohort study. *BMJ* 340:b4838
- Norberg M, Eriksson JW, Lindahl B et al (2006) A combination of HbA_{1c}, fasting glucose and BMI is effective in screening for individuals at risk of future type 2 diabetes: OGTT is not needed. *J Intern Med* 260:263–271
- Selvin E, Steffes MW, Zhu H et al (2010) Glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults. *N Engl J Med* 362:800–811
- Cheng P, Neugaard B, Foulis P, Conlin PR (2011) Hemoglobin A_{1c} as a predictor of incident diabetes. *Diabetes Care* 34:610–615
- Wang W, Lee ET, Howard BV, Fabsitz RR, Devereux RB, Welty TK (2011) Fasting plasma glucose and hemoglobin A_{1c} in

- identifying and predicting diabetes: the Strong Heart Study. *Diabetes Care* 34:363–368
26. Schottker B, Raum E, Rothenbacher D, Muller H, Brenner H (2011) Prognostic value of haemoglobin A_{1c} and fasting plasma glucose for incident diabetes and implications for screening. *Eur J Epidemiol* 26:779–787
 27. Abdul-Ghani MA, Abdul-Ghani T, Muller G et al (2011) Role of glycated hemoglobin in the prediction of future risk of T2DM. *J Clin Endocrinol Metab* 96:2596–2600
 28. Sattar N (2012) Biomarkers for diabetes prediction, pathogenesis or pharmacotherapy guidance? Past, present and future possibilities. *Diabet Med* 29:5–13
 29. Kolberg JA, Jorgensen T, Gerwien RW et al (2009) Development of a type 2 diabetes risk model from a panel of serum biomarkers from the Inter99 cohort. *Diabetes Care* 32:1207–1212
 30. American Diabetes Association (2010) Diagnosis and classification of diabetes mellitus. *Diabetes Care* 33(Suppl 1):S62–S69
 31. Sacks DB, Arnold M, Bakris GL et al (2011) Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. *Diabetes Care* 34:e61–e99
 32. World Health Organization (2011) Use of glycated haemoglobin (HbA_{1c}) in the diagnosis of diabetes mellitus. WHO, Geneva
 33. Heianza Y, Arase Y, Fujihara K et al (2012) Longitudinal trajectories of HbA_{1c} and fasting plasma glucose levels during the development of type 2 diabetes: the Toranomon Hospital Health Management Center Study 7 (TOPICS 7). *Diabetes Care* 35:1050–1052
 34. Kashiwagi A, Kasuga M, Araki E et al Committee on the Standardization of Diabetes Mellitus-Related Laboratory Testing of Japan Diabetes Society (JDS) (2012) International clinical harmonization of glycated hemoglobin in Japan: from Japan Diabetes Society to National Glycohemoglobin Standardization Program values. *Diabetology International* 3:8–10
 35. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44:837–845
 36. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS (2008) Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 27:157–172, discussion 207–212
 37. von Eckardstein A, Schulte H, Assmann G (2000) Risk for diabetes mellitus in middle-aged Caucasian male participants of the PROCAM study: implications for the definition of impaired fasting glucose by the American Diabetes Association. Prospective Cardiovascular Munster. *J Clin Endocrinol Metab* 85:3101–3108
 38. Kanaya AM, Wassel Fyr CL, de Rekeneire N et al (2005) Predicting the development of diabetes in older adults: the derivation and validation of a prediction rule. *Diabetes Care* 28:404–408
 39. Mann DM, Bertoni AG, Shimbo D et al (2010) Comparative validity of 3 diabetes mellitus risk prediction scoring models in a multiethnic US cohort: the Multi-Ethnic Study of Atherosclerosis. *Am J Epidemiol* 171:980–988
 40. Lim NK, Park SH, Choi SJ, Lee KS, Park HY (2012) A risk score for predicting the incidence of type 2 diabetes in a middle-aged Korean cohort. *Circ J* 76:1904–1910
 41. Willi C, Bodenmann P, Ghali WA, Faris PD, Cornuz J (2007) Active smoking and the risk of type 2 diabetes: a systematic review and meta-analysis. *JAMA* 298:2654–2664
 42. Alsema M, Vistisen D, Heymans MW et al (2011) The Evaluation of Screening and Early Detection Strategies for Type 2 Diabetes and Impaired Glucose Tolerance (DETECT-2) update of the Finnish diabetes risk score for prediction of incident type 2 diabetes. *Diabetologia* 54:1004–1012
 43. Hsu WC, Boyko EJ, Fujimoto WY et al (2012) Pathophysiologic differences among Asians, native Hawaiians, and other Pacific Islanders and treatment implications. *Diabetes Care* 35:1189–1198
 44. Chen L, Magliano DJ, Zimmet PZ (2012) The worldwide epidemiology of type 2 diabetes mellitus—present and future perspectives. *Nat Rev Endocrinol* 8:228–236
 45. Herman WH, Cohen RM (2012) Racial and ethnic differences in the relationship between HbA_{1c} and blood glucose: implications for the diagnosis of diabetes. *J Clin Endocrinol Metab* 97:1067–1072
 46. International Expert Committee (2009) International Expert Committee report on the role of the A_{1c} assay in the diagnosis of diabetes. *Diabetes Care* 32:1327–1334
 47. Zhang X, Gregg EW, Williamson DF et al (2010) A_{1c} level and future risk of diabetes: a systematic review. *Diabetes Care* 33:1665–1673
 48. Chatterjee R, Narayan KM, Lipscomb J, Phillips LS (2010) Screening adults for pre-diabetes and diabetes may be cost-saving. *Diabetes Care* 33:1484–1490
 49. Kahn R, Alperin P, Eddy D et al (2010) Age at initiation and frequency of screening to detect type 2 diabetes: a cost-effectiveness analysis. *Lancet* 375:1365–1374
 50. World Health Organization (2003) Screening for type 2 diabetes: report of a World Health Organization and International Diabetes Federation Meeting. WHO, Geneva