

Original Paper

Prediction of Glucose Metabolism Disorder Risk Using a Machine Learning Algorithm: Pilot Study

Katsutoshi Maeta¹, BE; Yu Nishiyama¹, PhD; Kazutoshi Fujibayashi², MD, PhD; Toshiaki Gunji³, MD, PhD; Noriko Sasabe³, MD, PhD; Kimiko Iijima³, MD, PhD; Toshio Naito², MD, PhD

¹Faculty of Informatics and Engineering, The University of Electro-Communications, Tokyo, Japan

²Department of General Medicine, School of Medicine, Juntendo University, Tokyo, Japan

³Center for Preventive Medicine, NTT Medical Center Tokyo, Tokyo, Japan

Corresponding Author:

Kazutoshi Fujibayashi, MD, PhD

Department of General Medicine

School of Medicine

Juntendo University

3-1-3

Hongo, Bunkyo-Ku

Tokyo,

Japan

Phone: 81 5802 1190

Email: kfujiba@juntendo.ac.jp

Abstract

Background: A 75-g oral glucose tolerance test (OGTT) provides important information about glucose metabolism, although the test is expensive and invasive. Complete OGTT information, such as 1-hour and 2-hour postloading plasma glucose and immunoreactive insulin levels, may be useful for predicting the future risk of diabetes or glucose metabolism disorders (GMD), which includes both diabetes and prediabetes.

Objective: We trained several classification models for predicting the risk of developing diabetes or GMD using data from thousands of OGTTs and a machine learning technique (XGBoost). The receiver operating characteristic (ROC) curves and their area under the curve (AUC) values for the trained classification models are reported, along with the sensitivity and specificity determined by the cutoff values of the Youden index. We compared the performance of the machine learning techniques with logistic regressions (LR), which are traditionally used in medical research studies.

Methods: Data were collected from subjects who underwent multiple OGTTs during comprehensive check-up medical examinations conducted at a single facility in Tokyo, Japan, from May 2006 to April 2017. For each examination, a subject was diagnosed with diabetes or prediabetes according to the American Diabetes Association guidelines. Given the data, 2 studies were conducted: predicting the risk of developing diabetes (study 1) or GMD (study 2). For each study, to apply supervised machine learning methods, the required label data was prepared. If a subject was diagnosed with diabetes or GMD at least once during the period, then that subject's data obtained in previous trials were classified into the risk group ($y=1$). After data processing, 13,581 and 6760 OGTTs were analyzed for study 1 and study 2, respectively. For each study, a randomly chosen subset representing 80% of the data was used for training 9 classification models and the remaining 20% was used for evaluating the models. Three classification models, A to C, used XGBoost with various input variables, some including OGTT data. The other 6 classification models, D to I, used LR for comparison.

Results: For study 1, the AUC values ranged from 0.78 to 0.93. For study 2, the AUC values ranged from 0.63 to 0.78. The machine learning approach using XGBoost showed better performance compared with traditional LR methods. The AUC values increased when the full OGTT variables were included. In our analysis using a particular setting of input variables, XGBoost showed that the OGTT variables were more important than fasting plasma glucose or glycated hemoglobin.

Conclusions: A machine learning approach, XGBoost, showed better prediction accuracy compared with LR, suggesting that advanced machine learning methods are useful for detecting the early signs of diabetes or GMD. The prediction accuracy increased when all OGTT variables were added. This indicates that complete OGTT information is important for predicting the future risk of diabetes and GMD accurately.

KEYWORDS

diabetes; machine learning; 75-g oral glucose tolerance test; XGBoost

Introduction

The incidence of diabetes has been increasing for the last decade and is expected to continue to increase in the future [1-3]. At present, diabetes is diagnosed and predicted based on fasting plasma glucose (FPG), glycated hemoglobin (HbA_{1c}), and plasma glucose levels 2 hours after a 75-g oral glucose tolerance test (OGTT) [4]. In an OGTT, a patient is asked to ingest a glucose drink, and their plasma glucose (PG) levels and immunoreactive insulin (IRI) levels are measured before and at intervals after the glucose drink is consumed. Although OGTT provides important information regarding pathological conditions of glucose metabolism, many diabetes survey tools predict the risk of diabetes development based only on noninvasive information, such as self-administered questionnaires [5]. The combination of parameters used to diagnose diabetes helps to identify individuals with a high risk of developing diabetes in the future. Heianza et al [6] showed that the combination of HbA_{1c} and FPG is useful for finding patients with a high risk of developing diabetes. Fujibayashi et al [7] used HbA_{1c} values, FPG levels, and 2-hour PG to predict instances of high future risk of developing diabetes. Complete data, including 1-hour and 2-hour PG and IRI values obtained by OGTT, may improve the prediction accuracy for diabetes risk.

Previously, logistic regression (LR) analyses were used as initial screening tests [5,8-10]. Recently, studies have demonstrated new methods, including machine learning algorithms, big data mining approaches, and genomic information, for the improved screening and prediction of diabetes [11,12]. Machine learning methods using all relevant information from OGTTs may be able to more accurately predict the risk of developing diabetes and prediabetes. The goal of this study was to verify this hypothesis. To our knowledge, no previous study has predicted the development of diabetes using all of the information from OGTTs combined with machine learning.

We used XGBoost [13,14] for machine learning, an advanced algorithm known for obtaining the winning solutions in data competitions such as Kaggle. In addition, XGBoost has been applied to other medical fields [15-17]. Gao et al [15] compared model-based approaches (such as LRs) and model-free approaches (including using XGBoost) for the task of forecasting falls by patients with Parkinson disease. The authors reported that the model-free approach provided more reliable forecasting. Nishio et al [16] applied XGBoost and support vector machine methods to the computer-aided diagnosis of lung nodules. The authors reported that XGBoost was generally superior to support

vector machine methods. Qiao et al [17] applied XGBoost and recurrent neural networks to a task of emergency room visit prediction. The authors reported that the nonlinear models had better performance than linear models.

Methods

Ethics Statement

This study was conducted using data from comprehensive periodic medical examinations at the Center for Preventive Medicine, NTT Medical Center Tokyo, from May 2006 to April 2017. In Japan, employers are required by the Industrial Safety and Health Law to commission medical examinations once a year to ensure the health of their employees. The Center for Preventive Medicine has been contracted by a telecommunications company, Nippon Telegraph and Telephone Corporation (NTT), to provide periodic medical examinations to their employees to comply with this law. This program involves comprehensive periodic medical examinations as well as many services beyond those mandated by law. The data used in this study were collected as part of this general health check-up program at the center. We retrieved subject clinical data from an institutional database, although the examinations were not specifically intended to collect new data for our study. Our research plan was announced on the websites of both our facility and the Center for Preventive Medicine. All subjects were informed that the clinical data obtained by the program would be retrospectively analyzed and published. In addition, it was announced that subjects could withdraw from our research study at any time. The study protocol was approved by the ethical review board of Juntendo University (No. 2017114) and the institutional ethics committee at the Center for Preventive Medicine (No. 17-664).

Study Population

Most of the study subjects were volunteers from among the employees of NTT and their families. They were primarily healthy office workers ranging in age from 40 to 60 years, with more male subjects than females. Our investigation focused on subjects who underwent a 75-g OGTT at the center between May 2006 to April 2017. Subjects without serious diabetes or advanced renal failure were assessed regarding the status of their glucose metabolism using the OGTT.

A total of 20,458 OGTT trials were collected from 9906 subjects during the period at the center. Table 1 shows the distribution of subjects with the number of OGTT trials obtained for each during the period. Overall, 6437 subjects underwent OGTT only once, while 1 subject had 12 OGTTs.

Table 1. Distribution of subjects according to the number of oral glucose tolerance test trials.

Trials undergone, n	Subjects, n
1	6437
2	1157
3	736
4	459
5	331
6	251
7	172
8	143
9	93
10	81
11	45
12	1

Data Collection

The examinations were performed on 2 consecutive days. On the first day, each patient's weight and height were measured after the removal of shoes and heavy clothing, and blood pressure was measured with an automatic monitor with the person in the sitting position. In addition, serum samples were collected from each participant after overnight fasting and immediately subjected to biochemical analysis. The blood samples were also used to determine each subject's HbA_{1c} level, which was measured using high-performance liquid chromatography with an automatic analyzer. On the second day, the subjects underwent an OGTT. We obtained the subjects' FPG levels along with 1-hour and 2-hour postloading PG IRI levels during the OGTT. The Japan Diabetes Society (JDS) HbA_{1c} values were converted to National Glycohemoglobin Standardization Program values using the formula developed by the JDS [18]: $HbA_{1c} = [HbA_{1c}(JDS)(\%) \times 1.02 + 0.25(\%)]$. Insulin sensitivity was calculated with the insulin sensitivity index (ISI; composite) [19,20]: $ISI(\text{composite}) = [10,000 / \sqrt{(\text{FPG level}(\text{mg/dL}) \times \text{fasting IRI level}(\mu\text{U/mL}) \times 2\text{-hour PG level}(\text{mg/dl}) \times 2\text{-hour IRI level}(\mu\text{U/mL}))}]$. The sum of plasma glucose (SPG) is defined as $SPG = \text{FPG levels} + 1\text{-hour PG level} + 2\text{-hour PG level}$. The sum of immunoreactive insulin (S-IRI) is defined as $S\text{-IRI} = \text{fasting IRI level} + 1\text{-hour IRI level} + 2\text{-hour IRI level}$.

We defined diabetes, normal glucose tolerance (NGT), and prediabetes according to the American Diabetes Association guidelines [4]. Diabetes is defined as subjects with an FPG level ≥ 126 mg/dL, a 2-hour postloading PG level ≥ 200 mg/dL, or an HbA_{1c} concentration $\geq 6.5\%$. NGT is defined as subjects with an FPG level < 100 mg/dL, a 2-hour postloading PG level < 140

mg/dL, and an HbA_{1c} level $< 5.7\%$. Prediabetes is defined as subjects without diabetes who failed to have NGT. In our study, we defined glucose metabolism disorders (GMD) as either diabetes or prediabetes.

Data Handling

Inclusion and Exclusion Flow

Initially, a total of 20,458 OGTT trials across all subjects were included. Data were removed based on the inclusion and exclusion criteria as shown in Figure 1. First, 6437 subjects who underwent OGTT only once during the period were excluded to increase the reliability of the data. Second, missing data were excluded, and the remaining number of trials was 14,020. Then, 2 studies were conducted: predicting the future risk of developing diabetes (study 1) or GMD (study 2).

Study 1. Prediction of Future Risk of Developing Diabetes

Study 1 was aimed at predicting the future risk of developing diabetes. To apply supervised machine learning to the data, data labels (at risk: $y=1$, not at risk: $y=0$) were required for each OGTT trial. It is widely known that diabetes can recur even after remission. In addition, women with a history of gestational diabetes have a high risk of developing diabetes in the future [21]. We considered that subjects with a diagnosis of diabetes in the past had a high risk of developing diabetes in the future. Because of this hypothesis, we defined the risk group as follows: a subject was in the risk group ($y=1$) for diabetes if he or she was diagnosed with diabetes at least once during the period. We defined a subject to be in the nonrisk group ($y=0$) if he or she did not belong to the risk group. From 14,020 trials, 439 data points from patients diagnosed with diabetes were excluded to focus only on nondiabetic subjects.

Figure 1. Inclusion and exclusion criteria.

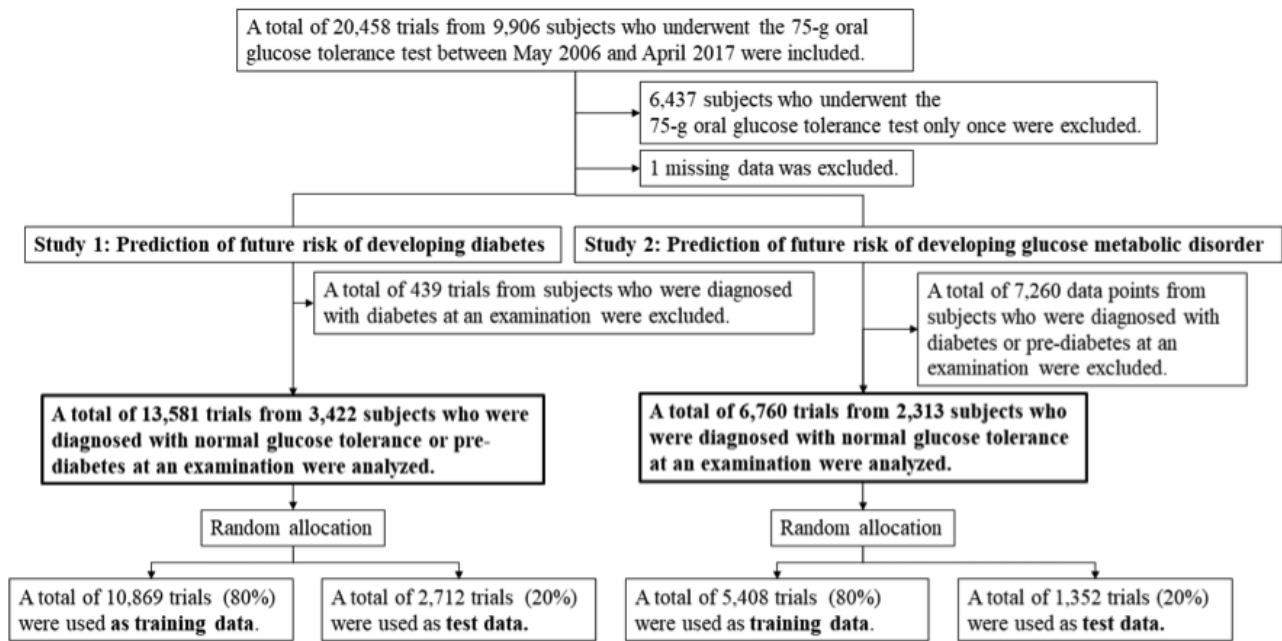


Figure 2. Examples of risk group and nonrisk group classifications. OGTT: oral glucose tolerance test, NGT: normal glucose tolerance.

Study 1			Study 2		
	Subject A	Subject B		Subject C	Subject D
First OGTT	NGT	Pre-diabetes	First OGTT	NGT	Pre-diabetes ^a
Second OGTT	Pre-diabetes	NGT	Second OGTT	NGT	NGT
Third OGTT	NGT	Diabetes ^a	Third OGTT	NGT	Diabetes ^a
Classification:	Non-risk group	Risk group	Classification:	Non-risk group	Risk group

Data with a superscript 'a' were excluded.

Examples of the risk group and nonrisk group for diabetes are shown in Figure 2 (left). Subjects A and B underwent OGTT 3 times. Subject B was diagnosed with prediabetes, NGT, and diabetes in the first, second, and third OGTTs, respectively. Thus, Subject B was classified into the risk group, since he or she was diagnosed with diabetes at least once during the period. The third OGTT data point, which occurred after the diagnosis (marked with superscript a), is removed to focus only on nondiabetic data. Subject A was diagnosed with NGT, prediabetes, and NGT in the first, second, and third OGTTs, respectively. Subject A was classified into the nonrisk group, since he or she was never diagnosed with diabetes during the period.

At the end, we had a total of 13,581 OGTT trials of patients who were diagnosed with NGT or prediabetes, each of which was labeled with future risk information (y=0 or y=1). We randomly selected 10,869 (80%) for the training data and used

the remaining 2712 (20%) for test data. A classification model was trained using the training data, and the prediction accuracy was evaluated with the test data. Nine classification models were trained, as described below. Table 2 shows a summary of the analyzed OGTT data. No significant differences were observed between the training and test data.

Study 2. Prediction of Future Risk of Glucose Metabolism Disorders

Study 2 was aimed at predicting the future risk of developing GMD, which includes either diabetes or prediabetes. Similar to study 1, we defined a subject as being in the risk group (y=1) for GMD if he or she was diagnosed with GMD (prediabetes or diabetes) at least once during the period. We defined a subject to be in the nonrisk group (y=0) if he or she did not belong to the risk group for GMD. From 14,020 trials, 7260 data points from patients diagnosed with GMD were excluded to focus only on NGT subjects.

Table 2. Statistical summary of the analyzed oral glucose tolerance test data (study 1).

Data points	Training data (n=10,869)	Test data (n=2712)	P value ^a
Age, years, mean (SD)	49.78 (9.27)	49.71 (9.16)	.71
Sex, male, n (%)	9998 (91.99)	2509 (92.51)	.61
Body mass index (kg/m ²), mean (SD)	23.43 (3.00)	23.35 (2.98)	.21
Glycated hemoglobin (%), mean (SD)	5.50 (0.32)	5.49 (0.31)	.21
FPG ^b (mg/dL), mean (SD)	96.59 (8.31)	96.47 (8.35)	.50
1-hour PG (mg/dL), mean (SD)	145.82 (40.67)	146.91 (40.55)	.21
2-hour PG (mg/dL), mean (SD)	111.07 (26.62)	111.26 (26.32)	.73
SPG ^c (mg/dL), mean (SD)	353.48 (63.99)	354.63 (63.51)	.40
Fasting IRI ^d (IU/mL), mean (SD)	6.41 (3.65)	6.40 (3.44)	.89
1-hour IRI (IU/mL), mean (SD)	55.50 (37.16)	55.04 (35.21)	.55
2-hour IRI (IU/mL), mean (SD)	41.95 (32.59)	41.32 (31.01)	.34
S-IRI ^e (IU/mL), mean (SD)	103.87 (64.95)	102.76 (61.65)	.41
ISI ^f (composite), mean (SD)	8.39 (5.27)	8.34 (4.89)	.59
Systolic BP ^g (mm Hg), mean (SD)	125.38 (17.65)	125.68 (17.37)	.42
Diastolic BP (mm Hg), mean (SD)	79.74 (11.22)	80.13 (11.21)	.10
Total cholesterol (mg/dL), mean (SD)	200.80 (30.95)	200.17 (31.01)	.35
HDL ^h (mg/dL), mean (SD)	58.27 (14.39)	58.10 (14.45)	.59
LDL ⁱ (mg/dL), mean (SD)	118.37 (28.12)	117.83 (28.40)	.38
Triglyceride (mg/dL), mean (SD)	115.07 (79.98)	114.07 (73.87)	.54
Uric acid (mg/dL), mean (SD)	6.20 (1.27)	6.17 (1.30)	.24
UN ^j (mg/dL), mean (SD)	13.70 (3.18)	13.57 (3.14)	.05
Serum creatinine (mg/dL), mean (SD)	0.88 (0.14)	0.88 (0.15)	.29
GOT ^k (IU/L), mean (SD)	24.02 (9.24)	23.81 (9.28)	.29
GPT ^l (IU/L), mean (SD)	25.56 (16.44)	24.95 (15.43)	.07
γ -GTP ^m (IU/L), mean (SD)	47.05 (53.13)	46.55 (48.97)	.64
Serum albumin (g/dL), mean (SD)	4.51 (0.26)	4.50 (0.26)	.11

^aUsed *t* test or chi-square test.

^bFPG: fasting plasma glucose.

^cSPG: sum of plasma glucose.

^dIRI: immunoreactive insulin.

^eS-IRI: sum of immunoreactive insulin.

^fISI: insulin sensitivity index.

^gBP: blood pressure.

^hHDL: high-density lipoprotein cholesterol.

ⁱLDL: low-density lipoprotein cholesterol.

^jUN: serum urea nitrogen.

^kGOT: serum glutamic oxaloacetic transaminase.

^lGPT: serum glutamic pyruvic transaminase.

^m γ -GTP: serum γ -glutamyl transpeptidase.

Figure 2 (right) shows examples of the risk group and nonrisk group for GMD. Subjects C and D underwent OGTT 3 times. Subject D was diagnosed with prediabetes, NGT, and diabetes

in the first, second, and third OGTTs, respectively. Subject D was classified into the risk group of GMD, since he or she was diagnosed with prediabetes at least once during the period. The

first and third OGTT, which were diagnosed with prediabetes and diabetes, respectively, were removed to focus only on NGT data. Subject C was diagnosed with NGT all 3 times. Therefore, subject C was classified into the nonrisk group, since he or she was never diagnosed with prediabetes or diabetes during the period.

Finally, we had 6760 OGTT trials of patients who were diagnosed with NGT, each of which was labeled with future risk information ($y=0$ or $y=1$). We randomly selected 5408 (80%) for training data, and used the remaining 1352 (20%) for test data. A classification model was trained using the training data, and the prediction accuracy was evaluated using the test data. Nine models were trained, as detailed in the Classification Models section. Table 3 shows a summary of the analyzed data. No significant differences were observed between the training and test data.

Statistical Analysis

XGBoost

XGBoost [13] is open-source software [14] that provides a machine learning method of regression and classification using ensemble learning with gradient tree boosting (GTB) [22]. XGBoost is well known for obtaining the winning solutions in data competitions. Chen and Guestrin [13] reported that “Among the 29 challenge-winning solutions published on Kaggle’s blog during 2015, 17 winning solutions used XGBoost.” Applications of XGBoost include practical tasks such as “store sales prediction, high energy physics event classification, Web text classification, customer behavior prediction, motion detection, ad click-through rate prediction, malware classification, product categorization, hazard risk prediction, and massive on-line course dropout rate prediction.” See Chen and Guestrin [13] for details of the applications. In addition, XGBoost has been applied to medical fields [15-17].

XGBoost (or GTB) learns a regression and classification function in the data space by sequentially optimizing weak learners, called regression trees. The parameters of a regression tree consist of the tree structures and the weights of the leaf nodes. They are sequentially optimized to minimize an objective function, consisting of a fitting loss term plus a regularization term, using gradient methods. XGBoost software is designed to increase the scalability and acceleration of optimized computation for practical use. See Chen and Guestrin [13] for technical details. The underlying GTB algorithm is briefly discussed in Multimedia Appendix 1. XGBoost includes several hyperparameters—including the maximum depth of regression trees, number of weak learners, learning rate, and regularization parameters—which need to be tuned.

Classification Models

To predict the future risk of diabetes (study 1) or GMD (study 2), we developed 9 classification models (A to I) with different input variables, shown in Table 4.

Models A, B, and C used XGBoost. For comparison, models D to I used LR, which is traditionally used in medical research studies. For each classification model, the input variables were set as follows. Model A inputs some basic variables relevant to diabetes or GMD, without OGTT variables. Model B inputs OGTT variables (1-hour PG, 1-hour IRI, 2-hour PG, and 2-hour IRI), as well as the variables of model A. Model C inputs all the measured variables. Blood pressure, lipid parameters, uric acid values, markers of liver function, and markers of kidney function are parameters related with metabolic syndrome, fatty liver, and chronic kidney disease. These conditions are associated with diabetes and were included as variables [23-25]. Models D to F served as baselines using the well-known biomarkers FPG and HbA_{1c} . Models G to I used the same variables as models A to C to directly compare the performances of XGBoost and LR.

Evaluation

To evaluate the 9 trained classification models, we used the receiver operating characteristic (ROC) curves and their area under the curve (AUC) values computed from the test data [26,27]. ROC curves have commonly been used in diabetes prediction research. In addition, we report the sensitivity and specificity at the cutoff values determined by the Youden index.

Hyperparameter Tuning of XGBoost

As mentioned, XGBoost includes several hyperparameters such as maximum depth of the regression trees, number of weak learners, learning rate, and regularization parameters that need to be tuned.

We tuned the parameters using a grid search to maximize the mean AUC value computed from 5-fold cross validation on the training data. Specifically, the training data were divided into 5 subsets at random: 4 subsets were used for training XGBoost and the other subset was used for validation. The ROC curve and AUC value can be evaluated from the validation subset. This procedure was repeated 5 times with different validation subsets. The mean AUC value can be computed by averaging the 5 AUC values. We tuned the hyperparameters, including the regularization parameters, with a grid search method to maximize the mean AUC value. After finding the optimal values of the hyperparameters, XGBoost was trained using the entire training data set. The final ROC and AUC values were then evaluated with the test data.

Given a ROC curve, a cutoff value is required to compute the sensitivity and specificity. The cutoff value was determined by averaging 5 cutoff values computed from the Youden index from 5-fold cross validation.

Table 3. Statistical summary of analyzed oral glucose tolerance test data (study 2).

Data points	Training data (n=5408)	Test data (n=1352)	P value ^a
Age, years, mean (SD)	47.13 (8.75)	46.93 (8.70)	.43
Sex, male, n (%)	4853 (89.74)	1210 (89.50)	.83
Body mass index (kg/m ²), mean (SD)	22.71 (2.70)	22.72 (2.70)	.89
Glycated hemoglobin (%), mean (SD)	5.32 (0.23)	5.31 (0.23)	.43
FPG ^b (mg/dL), mean (SD)	91.46 (4.99)	91.41 (5.01)	.74
1-hour PG (mg/dL), mean (SD)	129.47 (33.10)	128.86 (31.48)	.53
2-hour PG (mg/dL), mean (SD)	99.89 (18.41)	99.73 (18.34)	.77
SPG ^c (mg/dL), mean (SD)	320.81 (45.05)	320.00 (43.52)	.54
Fasting IRI ^d (IU/mL), mean (SD)	5.66 (3.13)	5.59 (2.71)	.47
1-hour IRI (IU/mL), mean (SD)	51.06 (33.02)	49.20 (31.85)	.06
2-hour IRI (IU/mL), mean (SD)	34.16 (22.88)	33.45 (21.68)	.28
S-IRI ^e (IU/mL), mean (SD)	90.88 (51.66)	88.24 (49.38)	.08
ISI ^f (composite), mean (SD)	9.93 (5.50)	9.98 (5.35)	.77
Systolic BP ^g (mm Hg), mean (SD)	121.77 (16.63)	121.97 (16.01)	.68
Diastolic BP (mm Hg), mean (SD)	77.55 (10.82)	77.38 (10.51)	.59
Total cholesterol (mg/dL), mean (SD)	196.35 (30.13)	196.35 (30.28)	>.99
HDL ^h (mg/dL), mean (SD)	59.53 (14.30)	60.25 (15.01)	.11
LDL ⁱ (mg/dL), mean (SD)	114.88 (27.66)	113.94 (27.93)	.27
Triglyceride (mg/dL), mean (SD)	101.64 (63.13)	102.36 (77.61)	.75
Uric acid (mg/dL), mean (SD)	6.05 (1.28)	6.06 (1.31)	.70
UN ^j (mg/dL), mean (SD)	13.35 (3.04)	13.43 (2.99)	.36
Serum creatinine (mg/dL), mean (SD)	0.87 (0.14)	0.87 (0.14)	.95
GOT ^k (IU/L), mean (SD)	22.96 (8.20)	22.81 (8.42)	.55
GPT ^l (IU/L), mean (SD)	23.12 (13.91)	22.72 (13.13)	.32
γ -GTP ^m (IU/L), mean (SD)	40.74 (45.54)	39.87 (46.02)	.53
Serum albumin (g/dL), mean (SD)	4.50 (0.26)	4.49 (0.27)	.47

^aUsed *t* test or chi-square test.

^bFPG: fasting plasma glucose.

^cSPG: sum of plasma glucose.

^dIRI: immunoreactive insulin.

^eS-IRI: sum of immunoreactive insulin.

^fISI: insulin sensitivity index.

^gBP: blood pressure.

^hHDL: high-density lipoprotein cholesterol.

ⁱLDL: low-density lipoprotein cholesterol.

^jUN: serum urea nitrogen.

^kGOT: serum glutamic oxaloacetic transaminase.

^lGPT: serum glutamic pyruvic transaminase.

^m γ -GTP: serum γ -glutamyl transpeptidase.

Table 4. List of trained classification models.

Model	Algorithm	Input variables
A	XGBoost	Sex, age, BMI ^a , HbA _{1c} ^b , FPG ^c , and fasting IRI ^d
B	XGBoost	Variables in model A, 1-hour PG, 2-hour PG, 1-hour IRI, and 2-hour IRI
C	XGBoost	Variables in model B, SPG ^e during the 75-g OGTT ^f , S-IRI ^g during the OGTT, simple ISI ^h (composite), systolic blood pressure, diastolic blood pressure, total cholesterol, HDLC ⁱ , LDLC ^j , triglyceride, uric acid, UN ^k , serum creatinine, GOT ^l , GPT ^m , γ -GPT ⁿ , and serum albumin
D	LR ^o	FPG
E	LR	HbA _{1c}
F	LR	FPG, HbA _{1c}
G	LR	Sex, age, BMI, HbA _{1c} , FPG, and fasting IRI
H	LR	Variables in model A, 1-hour PG, 2-hour PG, 1-hour IRI, and 2-hour IRI
I	LR	Variables in model B, SPG during the 75-g OGTT, S-IRI during the OGTT, simple ISI (composite), systolic blood pressure, diastolic blood pressure, total cholesterol, HDLC, LDLC, triglyceride, uric acid, UN, serum creatinine, GOT, GPT, γ -GPT, and serum albumin

^aBMI: body mass index.

^bHbA_{1c}: glycated hemoglobin.

^cFPG: fasting plasma glucose.

^dIRI: immunoreactive insulin.

^eSPG: sum of plasma glucose.

^fOGTT: oral glucose tolerance test.

^gS-IRI: sum of immunoreactive insulin.

^hISI: insulin sensitivity index.

ⁱHDLC: high-density lipoprotein cholesterol.

^jLDLC: low-density lipoprotein cholesterol.

^kUN: serum urea nitrogen.

^lGOT: serum glutamic oxaloacetic transaminase.

^mGPT: serum glutamic pyruvic transaminase.

ⁿ γ -GPT: serum γ -glutamyl transpeptidase.

^oLR: logistic regression.

Results

Study 1. Prediction of Future Risk of Developing Diabetes

Figure 3 shows the 6 ROC curves for models A to I. Similarly, Multimedia Appendix 2 shows the corresponding curves for models A to F. The horizontal axis represents the false positive rate, and the vertical axis represents the true positive rate. The 3 solid lines (models A, B, and C) show the ROC curves obtained from XGBoost. The dashed lines (models G, H, and I) show the ROC curves obtained from LR.

The AUC values for the 9 classification models are shown in Table 5. For each model, we also show the sensitivity and specificity as determined by the Youden index. We observed that XGBoost had superior performance compared with LR. The AUC value increased with the number of input variables. Models B and C, which exploit XGBoost and complete OGTT information for input variables, showed the best AUC values, 0.90 and 0.93, respectively.

In addition, XGBoost provides an importance score for each input variable. The importance value for each input variable in models A to C are shown in Multimedia Appendix 3 (left), Multimedia Appendix 4 (left), and Multimedia Appendix 5 (left), respectively. In model B, we observed that the OGTT variables (1-hour PG, 1-hour IRI, 2-hour PG, and 2-hour IRI) were more important than FPG or HbA_{1c}. In model C, we observed that SPG and 2-hour PG were more important variables than FPG or HbA_{1c}, although multicollinearity needs to be considered.

Study 2. Prediction of Future Risk of Glucose Metabolism Disorders

Figure 4 shows the 6 ROC curves for models A to I. Similarly, Multimedia Appendix 6 shows the corresponding 6 ROC curves for models A to F. The 3 solid lines (models A, B, and C) show the ROC curves obtained from XGBoost. The dashed lines (models G, H, and I) show the ROC curves obtained from LR.

Figure 3. Receiver operating characteristic curves obtained for the prediction of diabetes.

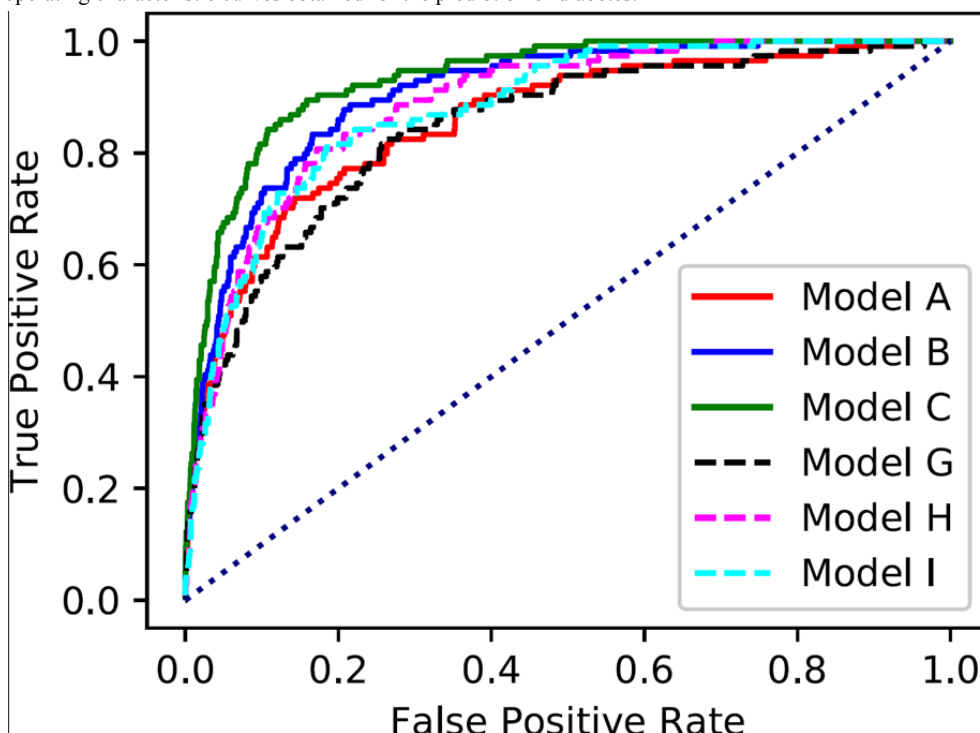


Table 5. Area under the curve, sensitivity, and specificity for predicting diabetes.

Type and model	AUC ^a	Sensitivity (%)	Specificity (%)
Machine learning			
Model A	0.86	36.8	97.2
Model B	0.90	40.4	97.4
Model C	0.93	39.5	98.4
Logistic regression			
Model D	0.80	12.3	99.4
Model E	0.78	6.1	99.9
Model F	0.84	41.2	95.2
Model G	0.85	38.6	95.9
Model H	0.88	37.7	96.1
Model I	0.88	40.4	96.5

^aAUC: area under the curve.

The AUC values for the 9 models are shown in Table 6. The sensitivity and specificity, as determined by the Youden index, are also shown in Table 6. Similar to study 1, we observed that XGBoost had better performance than LR. The AUC values also increased with the number of input variables. Models B and C, which exploit XGBoost and complete OGTT information as input variables, displayed the highest AUC values, 0.75 and 0.78, respectively.

The importance score of each input variable for models A, B, and C are shown in Multimedia Appendix 3 (right), Multimedia Appendix 4 (right), and Multimedia Appendix 5 (right), respectively. In model B, we observed that the OGTT variables (1-hour PG, 1-hour IRI, 2-hour PG, and 1-hour IRI) were more important than FPG or HbA_{1c}. In model C, we observed that the OGTT variables (1-hour PG, 1-hour IRI, 2-hour PG, 2-hour IRI, and SPG) were more important variables than FPG or HbA_{1c}, although multicollinearity needs to be considered.

Figure 4. Receiver operating characteristic curves for the prediction of glucose metabolism disorders.

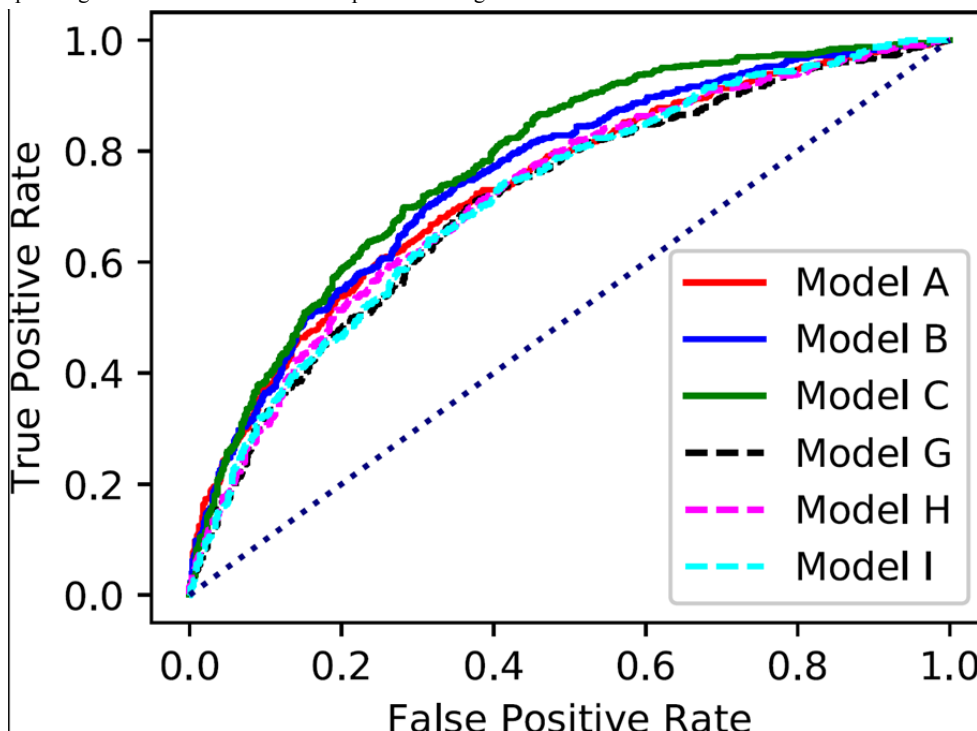


Table 6. Area under the curve, sensitivity, and specificity for predicting glucose metabolism disorders.

Type and model	AUC ^a	Sensitivity (%)	Specificity (%)
Machine learning			
Model A	0.73	34.3	90.9
Model B	0.75	31.6	91.8
Model C	0.78	33.7	92.2
Logistic regression			
Model D	0.65	6.6	97.3
Model E	0.63	21.3	91.6
Model F	0.69	6.6	98.4
Model G	0.71	27.6	78.9
Model H	0.72	27.1	91.6
Model I	0.72	41.3	84.9

^aAUC: area under the curve.

Discussion

Principal Findings

In this study, we reported on 2 results for predicting the future risk of diabetes or GMD using complete OGTT information and machine learning.

A feature of our study is that we used a large-scale dataset that included thousands of Japanese OGTT trials, even though OGTT is expensive and invasive. The amount of data enabled us to use a machine learning approach. It is known that one of the earliest detectable abnormalities in the development of diabetes is the deterioration of the early insulin response after glucose loading [28], and the aggravation of insulin sensitivity affects

the development of diabetes [7,29-32]. Data obtained from OGTT provide important pathological glucose metabolism information. We believe that the data obtained from OGTT contributed to the improvement in prediction of future risk of diabetes and GMD.

Another feature of our study is that we used an advanced and powerful machine learning method, XGBoost, which resulted in better performance compared with LR. Many previous diabetes risk assessment tools used LR analyses. Recently, various machine learning algorithms have been used for the screening and prediction of diabetes [12,33-35]. Linear approaches are generally unsuited for prediction models with complex correlations. We believe that XGBoost plays an important role in improving the prediction of the future risk of

developing diabetes or GMD. To our knowledge, no previous study combined large-scale Japanese OGTT data and XGBoost.

By observing the importance scores ([Multimedia Appendix 3](#)-[Multimedia Appendix 5](#)) of the input variables computed from XGboost, the OGTT variables (1-hour PG, 1-hour IRI, 2-hour PG, 2-hour IRI) were found to be more important predictors than FPG or HbA_{1c} for the future risk of diabetes or GMD, although multicollinearity needs to be taken into account. The progress of the PG level after loading can reflect abnormalities in the insulin response. Simultaneous measurement of PG and IRI levels can help evaluate the insulin sensitivity. Although multiple collinearity affects the results, PG and IRI levels after loading appeared to be more important than FPG or HbA_{1c}.

Limitations

Our research had several limitations. First, we did not use any information obtained from questionnaires in our research. This was because we were concerned that subject recall bias may impact the accuracy of the predictions [36]. As far as we know, previous diabetes risk assessment tools were based on information obtained from questionnaires such as family history and lifestyle habits [5,8-10,37]. Because of this, our results could not be easily compared with these reports. A previous study showed that combining the results of blood tests and questionnaire information improved the prediction accuracy of diabetes risk assessment [37]. We will attempt to improve the accuracy of diabetes risk prediction by integrating information obtained from blood tests and questionnaires. Second, we merged data from subjects who underwent OGTT different numbers of times. That is, we handled data from subjects who had 2 OGTT trials in the same manner as subjects with 10 OGTT trials. Data from subjects who had frequent OGTT trials may have impacted the calculation. Finally, our subjects were affected by selection bias, specifically, the “healthy worker”

effect. More than 70% of our participants were healthy male office workers who ranged in age from 40 to 60 years. Thus, the limited sample might not accurately represent the entire population. In addition, we believe that our method is not suitable for predicting rapidly progressing diabetes, as with type 1 diabetes. Also, validation is still required using other data sets.

Comparison With Prior Work

Thoopputra et al [5] considered many diabetes risk assessment tools developed worldwide. In the review, although there a few that used OGTT data or decision tree algorithms, many diabetes risk assessment tools used only noninvasive information and LR analyses. Values for the AUC ranged from 62% to 87%. In Japan, Nanri et al [37] reported a risk score showing an AUC value of 0.882 for predicting type 2 diabetes based on noninvasive information, FPG level and HbA_{1c}, using an LR analysis. Recently, studies have demonstrated new methods, including machine learning algorithms, big data mining approaches, and genomic information, for the screening and prediction of diabetes [11,12]. Habibi et al [34] developed a model with an AUC value of 0.875 when screening for type 2 diabetes that used a decision tree method and did not require any laboratory tests. López et al [12] reported a model for diabetes prediction having an AUC value of 0.89 that used genetic information and a random forest algorithm.

Conclusion

Our predictions for the future risk of developing diabetes or GMD, using data from thousands of OGTT trials and the machine learning program XGBoost, resulted in higher accuracy compared with traditional LR analysis. Combining complete OGTT information with advanced machine learning algorithms may be useful for detecting the future risk of diabetes or GMD more accurately.

Acknowledgments

We thank Louis R Nemzer, PhD, from Edanz Group for editing a draft of this manuscript.

Conflicts of Interest

This study was funded by Novartis Pharma KK.

Multimedia Appendix 1

Brief review of gradient tree boosting algorithms.

[[PDF File \(Adobe PDF File\), 57KB - diabetes_v3i4e10212_app1.pdf](#)]

Multimedia Appendix 2

Receiver operating characteristic curves in study 1, comparison with reference models D-F.

[[PPTX File, 237KB - diabetes_v3i4e10212_app2.pptx](#)]

Multimedia Appendix 3

Importance of variables in model A.

[[PPTX File, 58KB - diabetes_v3i4e10212_app3.pptx](#)]

Multimedia Appendix 4

Importance of variables in model B.

[[PPTX File, 68KB - diabetes_v3i4e10212_app4.pptx](#)]

Multimedia Appendix 5

Importance of variables in model C.

[[PPTX File, 124KB - diabetes_v3i4e10212_app5.pptx](#)]

Multimedia Appendix 6

Receiver operating characteristic curves for study 2, comparisons with reference models D-F.

[[PPTX File, 240KB - diabetes_v3i4e10212_app6.pptx](#)]

References

1. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Res Clin Pract* 2010 Jan;87(1):4-14. [doi: [10.1016/j.diabres.2009.10.007](#)] [Medline: [19896746](#)]
2. Whiting DR, Guariguata L, Weil C, Shaw J. IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. *Diabetes Res Clin Pract* 2011 Dec;94(3):311-321. [doi: [10.1016/j.diabres.2011.10.029](#)] [Medline: [22079683](#)]
3. Ogurtsova K, da Rocha Fernandes JD, Huang Y, Linnenkamp U, Guariguata L, Cho NH, et al. IDF Diabetes Atlas: global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes Res Clin Pract* 2017 Jun;128:40-50. [doi: [10.1016/j.diabres.2017.03.024](#)] [Medline: [28437734](#)]
4. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2010 Jan;33 Suppl 1:S62-S69 [FREE Full text] [doi: [10.2337/dc10-S062](#)] [Medline: [20042775](#)]
5. Thooputra T, Newby D, Schneider J, Li SC. Survey of diabetes risk assessment tools: concepts, structure and performance. *Diabetes Metab Res Rev* 2012 Sep;28(6):485-498. [doi: [10.1002/dmrr.2296](#)] [Medline: [22407958](#)]
6. Heianza Y, Hara S, Arase Y, Saito K, Fujiwara K, Tsuji H, et al. HbA1c 5.7-6.4% and impaired fasting plasma glucose for diagnosis of prediabetes and risk of progression to diabetes in Japan (TOPICS 3): a longitudinal cohort study. *Lancet* 2011 Jul 09;378(9786):147-155. [doi: [10.1016/S0140-6736\(11\)60472-8](#)] [Medline: [21705064](#)]
7. Fujibayashi K, Yokokawa H, Gunji T, Sasabe N, Okumura M, Iijima K, et al. Utility of 75-g oral glucose tolerance test results and hemoglobin a1c values for predicting the incidence of diabetes mellitus among middle-aged Japanese men: a large-scale retrospective cohort study performed at a single hospital. *Intern Med* 2015;54(7):717-723 [FREE Full text] [doi: [10.2169/internalmedicine.54.2839](#)] [Medline: [25832931](#)]
8. Fujiati II, Damanik HA, Bachtiar A, Nurdin AA, Ward P. Development and validation of prediabetes risk score for predicting prediabetes among Indonesian adults in primary care: cross-sectional diagnostic study. *Interv Med Appl Sci* 2017 Jun;9(2):76-85 [FREE Full text] [doi: [10.1556/1646.9.2017.2.18](#)] [Medline: [28932501](#)]
9. Heikes KE, Eddy DM, Arondekar B, Schlessinger L. Diabetes Risk Calculator: a simple tool for detecting undiagnosed diabetes and pre-diabetes. *Diabetes Care* 2008 May;31(5):1040-1045. [doi: [10.2337/dc07-1150](#)] [Medline: [18070993](#)]
10. Heltberg A, Andersen JS, Sandholdt H, Siersma V, Kragstrup J, Ellervik C. Predictors of undiagnosed prevalent type 2 diabetes: the Danish General Suburban Population Study. *Prim Care Diabetes* 2018 Dec;12(1):13-22. [doi: [10.1016/j.pcd.2017.08.005](#)] [Medline: [28964672](#)]
11. Kavakiotis I, Tsave O, Salifoglou A, Maglaveras N, Vlahavas I, Chouvarda I. Machine learning and data mining methods in diabetes research. *Comput Struct Biotechnol J* 2017;15:104-116 [FREE Full text] [doi: [10.1016/j.csbj.2016.12.005](#)] [Medline: [28138367](#)]
12. López B, Torrent-Fontbona F, Viñas R, Fernández-Real JM. Single Nucleotide Polymorphism relevance learning with Random Forests for Type 2 diabetes risk prediction. *Artif Intell Med* 2018 Dec;85:43-49. [doi: [10.1016/j.artmed.2017.09.005](#)] [Medline: [28943335](#)]
13. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. 2016 Presented at: KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016; San Francisco p. 785-794.
14. XGBoost. URL: <https://github.com/dmlc/xgboost> [accessed 2018-02-22] [WebCite Cache ID 6xQqVe6mN]
15. Gao C, Sun H, Wang T, Tang M, Bohnen NI, Müller MLTM, et al. Model-based and model-free machine learning techniques for diagnostic prediction and classification of clinical outcomes in parkinson's disease. *Sci Rep* 2018 May 08;8(1):7129 [FREE Full text] [doi: [10.1038/s41598-018-24783-4](#)] [Medline: [29740058](#)]
16. Nishio M, Nishizawa M, Sugiyama O, Kojima R, Yakami M, Kuroda T, et al. Computer-aided diagnosis of lung nodule using gradient tree boosting and Bayesian optimization. *PLoS One* 2018;13(4):e0195875 [FREE Full text] [doi: [10.1371/journal.pone.0195875](#)] [Medline: [29672639](#)]
17. Qiao Z, Sun N, Li X, Xia E, Zhao S, Qin Y. Using machine learning approaches for emergency room visit prediction based on electronic health record data. *Stud Health Technol Inform* 2018;247:111-115. [Medline: [29677933](#)]

18. Kashiwagi A, Kasuga M, Araki E, Oka Y, Hanafusa T, Ito H, Committee on the Standardization of Diabetes Mellitus-Related Laboratory Testing of Japan Diabetes Society. International clinical harmonization of glycated hemoglobin in Japan: from Japan Diabetes Society to National Glycohemoglobin Standardization Program values. *J Diabetes Investig* 2012 Feb 20;3(1):39-40 [FREE Full text] [doi: [10.1111/j.2040-1124.2012.00207.x](https://doi.org/10.1111/j.2040-1124.2012.00207.x)] [Medline: [24843544](https://pubmed.ncbi.nlm.nih.gov/24843544/)]
19. Matsuda M, DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. *Diabetes Care* 1999 Sep;22(9):1462-1470 [FREE Full text] [Medline: [10480510](https://pubmed.ncbi.nlm.nih.gov/10480510/)]
20. DeFronzo RA, Matsuda M. Reduced time points to calculate the composite index. *Diabetes Care* 2010 Jul;33(7):e93. [doi: [10.2337/dc10-0646](https://doi.org/10.2337/dc10-0646)] [Medline: [20587713](https://pubmed.ncbi.nlm.nih.gov/20587713/)]
21. Bellamy L, Casas J, Hingorani AD, Williams D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. *Lancet* 2009 May 23;373(9677):1773-1779. [doi: [10.1016/S0140-6736\(09\)60731-5](https://doi.org/10.1016/S0140-6736(09)60731-5)] [Medline: [19465232](https://pubmed.ncbi.nlm.nih.gov/19465232/)]
22. Friedman J. Greedy function approximation: a gradient boosting machine. *Ann Statist* 2001;29(5):1189-1232.
23. Samson SL, Garber AJ. Metabolic syndrome. *Endocrinol Metab Clin North Am* 2014 Mar;43(1):1-23. [doi: [10.1016/j.ecl.2013.09.009](https://doi.org/10.1016/j.ecl.2013.09.009)] [Medline: [24582089](https://pubmed.ncbi.nlm.nih.gov/24582089/)]
24. Leite NC, Villela-Nogueira CA, Cardoso CRL, Salles GF. Non-alcoholic fatty liver disease and diabetes: from physiopathological interplay to diagnosis and treatment. *World J Gastroenterol* 2014 Jul 14;20(26):8377-8392 [FREE Full text] [doi: [10.3748/wjg.v20.i26.8377](https://doi.org/10.3748/wjg.v20.i26.8377)] [Medline: [25024596](https://pubmed.ncbi.nlm.nih.gov/25024596/)]
25. Mantovani A, Zaza G, Byrne CD, Lonardo A, Zoppini G, Bonora E, et al. Nonalcoholic fatty liver disease increases risk of incident chronic kidney disease: a systematic review and meta-analysis. *Metabolism* 2018 Dec;79:64-76. [doi: [10.1016/j.metabol.2017.11.003](https://doi.org/10.1016/j.metabol.2017.11.003)] [Medline: [29137912](https://pubmed.ncbi.nlm.nih.gov/29137912/)]
26. Metz CE. Receiver operating characteristic analysis: a tool for the quantitative evaluation of observer performance and imaging systems. *J Am Coll Radiol* 2006 Jun;3(6):413-422. [doi: [10.1016/j.jacr.2006.02.021](https://doi.org/10.1016/j.jacr.2006.02.021)] [Medline: [17412096](https://pubmed.ncbi.nlm.nih.gov/17412096/)]
27. Krzanowski W, Hand DJ. ROC Curves for Continuous Data (Chapman & Hall/CRC Monographs on Statistics & Applied Probability). Boca Raton: Chapman and Hall; 2009.
28. Gerich JE. Is reduced first-phase insulin release the earliest detectable abnormality in individuals destined to develop type 2 diabetes? *Diabetes* 2002 Feb;51 Suppl 1:S117-S121 [FREE Full text] [Medline: [11815469](https://pubmed.ncbi.nlm.nih.gov/11815469/)]
29. Osei K, Rhinesmith S, Gaillard T, Schuster D. Impaired insulin sensitivity, insulin secretion, and glucose effectiveness predict future development of impaired glucose tolerance and type 2 diabetes in pre-diabetic African Americans: implications for primary diabetes prevention. *Diabetes Care* 2004 Jun;27(6):1439-1446. [Medline: [15161801](https://pubmed.ncbi.nlm.nih.gov/15161801/)]
30. Abdul-Ghani MA, Williams K, DeFronzo R, Stern M. Risk of progression to type 2 diabetes based on relationship between postload plasma glucose and fasting plasma glucose. *Diabetes Care* 2006 Jul;29(7):1613-1618. [doi: [10.2337/dc05-1711](https://doi.org/10.2337/dc05-1711)] [Medline: [16801587](https://pubmed.ncbi.nlm.nih.gov/16801587/)]
31. Abdul-Ghani MA, Williams K, DeFronzo RA, Stern M. What is the best predictor of future type 2 diabetes? *Diabetes Care* 2007 Jun;30(6):1544-1548. [doi: [10.2337/dc06-1331](https://doi.org/10.2337/dc06-1331)] [Medline: [17384342](https://pubmed.ncbi.nlm.nih.gov/17384342/)]
32. Oka R, Aizawa T, Miyamoto S, Yoneda T, Yamagishi M. One-hour plasma glucose as a predictor of the development of Type 2 diabetes in Japanese adults. *Diabet Med* 2016 Dec;33(10):1399-1405. [doi: [10.1111/dme.12994](https://doi.org/10.1111/dme.12994)] [Medline: [26482027](https://pubmed.ncbi.nlm.nih.gov/26482027/)]
33. Razavian N, Blecker S, Schmidt AM, Smith-McLallen A, Nigam S, Sontag D. Population-level prediction of type 2 diabetes from claims data and analysis of risk factors. *Big Data* 2015 Dec;3(4):277-287. [doi: [10.1089/big.2015.0020](https://doi.org/10.1089/big.2015.0020)] [Medline: [27441408](https://pubmed.ncbi.nlm.nih.gov/27441408/)]
34. Habibi S, Ahmadi M, Alizadeh S. Type 2 diabetes mellitus screening and risk factors using decision tree: results of data mining. *Glob J Health Sci* 2015 Mar 18;7(5):304-310 [FREE Full text] [doi: [10.5539/gjhs.v7n5p304](https://doi.org/10.5539/gjhs.v7n5p304)] [Medline: [26156928](https://pubmed.ncbi.nlm.nih.gov/26156928/)]
35. Anderson JP, Parikh JR, Shenfeld DK, Ivanov V, Marks C, Church BW, et al. Reverse engineering and evaluation of prediction models for progression to type 2 diabetes: an application of machine learning using electronic health records. *J Diabetes Sci Technol* 2015 Dec 20;10(1):6-18 [FREE Full text] [doi: [10.1177/1932296815620200](https://doi.org/10.1177/1932296815620200)] [Medline: [26685993](https://pubmed.ncbi.nlm.nih.gov/26685993/)]
36. Schmier JK, Halpern MT. Patient recall and recall bias of health state and health status. *Expert Rev Pharmacoecon Outcomes Res* 2004 Apr;4(2):159-163. [doi: [10.1586/14737167.4.2.159](https://doi.org/10.1586/14737167.4.2.159)] [Medline: [19807519](https://pubmed.ncbi.nlm.nih.gov/19807519/)]
37. Nanri A, Nakagawa T, Kuwahara K, Yamamoto S, Honda T, Okazaki H, Japan Epidemiology Collaboration on Occupational Health Study Group. Correction: development of risk score for predicting 3-year incidence of type 2 diabetes: Japan Epidemiology Collaboration on Occupational Health Study. *PLoS One* 2018;13(6):e0199075 [FREE Full text] [doi: [10.1371/journal.pone.0199075](https://doi.org/10.1371/journal.pone.0199075)] [Medline: [29879228](https://pubmed.ncbi.nlm.nih.gov/29879228/)]

Abbreviations

- AUC:** area under the curve
- BMI:** body mass index
- BP:** blood pressure
- FPG:** fasting plasma glucose
- GMD:** glucose metabolism disorders
- GOT:** serum glutamic oxaloacetic transaminase
- GPT:** serum glutamic pyruvic transaminase

GTB: gradient tree boosting
HbA_{1c}: glycated hemoglobin
HDLC: high-density lipoprotein cholesterol
IRI: immunoreactive insulin
ISI: insulin sensitivity index
JDS: Japan Diabetes Society
LDLC: low-density lipoprotein cholesterol
LR: logistic regression
NGT: normal glucose tolerance
NTT: Nippon Telegraph and Telephone Corporation
OGTT: oral glucose tolerance test
PG: plasma glucose
ROC: receiver operating characteristic
S-IRI: sum of immunoreactive insulin
SPG: sum of plasma glucose
UN: serum urea nitrogen
γ-GTP: serum γ-glutamyl transpeptidase

Edited by M Focsa; submitted 24.02.18; peer-reviewed by B Lopez, J Bian, A Billis, MF Kabir, L Li; comments to author 11.04.18; revised version received 16.08.18; accepted 17.10.18; published 26.11.18

Please cite as:

Maeta K, Nishiyama Y, Fujibayashi K, Gunji T, Sasabe N, Iijima K, Naito T
Prediction of Glucose Metabolism Disorder Risk Using a Machine Learning Algorithm: Pilot Study
JMIR Diabetes 2018;3(4):e10212
URL: <http://diabetes.jmir.org/2018/4/e10212/>
doi: [10.2196/10212](https://doi.org/10.2196/10212)
PMID: [30478026](https://pubmed.ncbi.nlm.nih.gov/30478026/)

©Katsutoshi Maeta, Yu Nishiyama, Kazutoshi Fujibayashi, Toshiaki Gunji, Noriko Sasabe, Kimiko Iijima, Toshio Naito. Originally published in *JMIR Diabetes* (<http://diabetes.jmir.org>), 26.11.2018. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Diabetes*, is properly cited. The complete bibliographic information, a link to the original publication on <http://diabetes.jmir.org/>, as well as this copyright and license information must be included.