Published in final edited form as:

Diabetes Res Clin Pract. 2007 October ; 78(1): 93–101. doi:10.1016/j.diabres.2007.02.013.

Individualized pharmacokinetic risk assessment for development of diabetes in high risk population

N. Gupta^a, N.H. Al-Huniti^b, and P. Veng-Pedersen^{c,*}

^a Abbott Laboratories, Abbott Park, IL, USA

^b Pharsight Corporation, Cary, NC, USA

^c College of Pharmacy, University of Iowa, Iowa City, IA 52242, USA

Abstract

Aim—The objective of this study is to propose a non-parametric pharmacokinetic prediction model that addresses the individualized risk of developing type-2 diabetes in subjects with family history of type-2 diabetes.

Method—All selected 191 healthy subjects had both parents as type-2 diabetic. Glucose was administered intravenously (0.5 g/kg body weight) and 13 blood samples taken at specified times were analyzed for plasma insulin and glucose concentrations. All subjects were followed for an average of 13–14 years for diabetic or normal (non-diabetic) outcome.

Results—The new logistic regression model predicts the development of diabetes based on body mass index and only one blood sample at 90 min analyzed for insulin concentration. Our model correctly identified 4.5 times more subjects (54% versus 11.6%) predicted to develop diabetes and more than twice the subjects (99% versus 46.4%) predicted not to develop diabetes compared to current non-pharmacokinetic probability estimates for development of type-2 diabetes.

Conclusion—Our model can be useful for individualized prediction of development of type-2 diabetes in subjects with family history of type-2 diabetes. This improved prediction may be an important mediating factor for better perception of risk and may result in an improved intervention.

Keywords

Pharmacokinetics; Family history; Predictive model; Type-2 diabetes

Undiagnosed diabetes constitutes approximately 29.3% of total diabetes prevalence [1]. It is clear that developing strategies to screen and identify high-risk individuals should be an important public health goal [2]. Family history of type-2 diabetes is recognized as an important risk factor of the disease [3–5]. Individuals who have a family history of diabetes have two to six times greater risk of developing type-2 diabetes compared to individuals with no family history of the disease [6]. According to NHANES III, from 1999–2002, prevalence of diabetes among individuals who have a first degree relative (parents or siblings) with diabetes (14.3%) was significantly higher ($p < 0.001$) than that of individuals without a family history of diabetes (3.2%) [2]. According to the same survey, the probability of developing diabetes among individuals who have both parents with diabetes (25.4%) was significantly higher ($p < 0.001$) than individuals without any family history of diabetes (3.2%) [2]. The diabetes prevalence in subjects with BMI < 25 kg/m² was 3.1% followed by 5.9% for subjects with BMI from 25 to

* Corresponding author. Tel.: +1 319 335 8792; fax: +1 319 335 9349. veng@uiowa.edu (P. Veng-Pedersen).

29.9 kg/m² and for subjects with BMI $\geq$ 30 kg/m², the prevalence of diabetes was 11.2% [2]. Thus the magnitude of increased risk is greater for family history than for obesity *per se*.

Although family history is a risk factor for diabetes, various studies have shown that fewer than 40% of people with a family history of the disease actually perceive themselves to be at an increased risk of developing diabetes compared to those with no family history of diabetes [7–11]. Therefore, altering risk perception is a potential target for intervention. This may be achieved by a more accurate and individualized risk assessment. An increased awareness of the risk of developing diabetes, will likely increase an individual's motivation for changing his or her lifestyle to reduce the risk of developing the disease [12,13]. There are no studies reported in the diabetes literature that specifically address a scientific-based manipulation of individualized risk perception in subjects with family history of type-2 diabetes, although studies in other areas have demonstrated a direct relationship between altering risk perception and changing human behavior [6]. In this study, we have proposed and validated a pharmacokinetic predictive model that addresses the individualized risk perception in subjects with family history of type-2 diabetes (both parents type-2 diabetic). It is known that type-2 diabetes is a heterogeneous disorder characterized by a combination of impaired insulin secretion and insulin resistance [14,15]. Our exploratory data analysis is aimed at identification of few plasma insulin concentrations that can improve the prediction of developing type-2 diabetes. This is an important goal because the use of only few (one or two) blood samples analyzed for insulin concentration makes the test considerably more practical than tests requiring many samples for a pharmacokinetic/pharmacodynamic characterization of the glucose–insulin system, such as in the intravenous glucose tolerance test (IVGTT).

Our model is useful for individualized prediction of development of type-2 diabetes in subjects with family history of type-2 diabetes. The improvement in prediction may be an important mediating factor for better perception of risk and may result in an improved intervention.

1. Materials and methods

1.1. Subjects

For this study 191 healthy subjects from age 16 to 59 years with family history of type-2 diabetes (i.e., both parents had type-2 diabetes) were selected [16,17]. At the time of recruitment, no subject had a disease or received any medication known to affect glucose metabolism. Follow-up survey of all the subjects with diabetic parents was conducted after an average of 13 years (12.7 ± 6.4 years) (mean $\pm$ SD) to ascertain diagnoses of diabetes. Verification of the absence of type-2 diabetes was based on an oral glucose tolerance test. Subjects were divided into two groups based on the outcome, diabetic and normal (non-diabetic) outcome. Subjects with missing information (i.e., demographics, outcome at the end of follow-up period, insulin or glucose concentrations during IVGTT) were not included in the analysis. We had 30 subjects in the diabetic outcome group (19.7%) and 122 subjects in the normal outcome group (80.3%). The body mass index (BMI) was significantly higher ($p < 0.01$) in the diabetic outcome group (31.3 ± 7.6 kg/m²) as compared to the normal outcome group (25.3 ± 4.3). The age at entry for subjects in diabetic group (36.3 ± 9.2 years) was not significantly different ($p > 0.05$) than subjects in normal outcome group (33.6 ± 9.9).

1.2. Intravenous glucose tolerance test

All the participants were instructed to consume a high-carbohydrate diet (250–300 g/day) for at least 3 days before the test and to arrive in the clinical research laboratory at Joslin Diabetes Center (Boston, MA) after an overnight fast. Height and weight were measured and a medical history was taken. After a fasting blood sample was obtained, glucose was administered intravenously (0.5 g/kg body weight), and blood was drawn at 1, 3, 5, 10, 20, 30, 40, 50, 60,

90, 120 and 180 min. Blood glucose concentration (mg/dL) was measured by either ferricyanide [18] or glucose oxidase method [19]. Serum insulin level ($\mu\text{U/mL}$) was measured by a double-antibody radioimmunoassay technique [20].

1.3. Data analysis

We combined subjects that developed diabetes and subjects that developed impaired glucose tolerance (IGT) in one group (called diabetic outcome group) and compared this group with the individuals that did not develop diabetes in the follow-up (normal outcome group). Subjects with IGT outcome were pooled with subjects with diabetic outcome because IGT is associated with higher risk of developing diabetes [21,22]. The risk of development of diabetes was 6.3 times higher in subjects with IGT than in normal subjects [21]. The criteria for diagnosis of IGT and diabetes is based on classification by National Diabetes Data Group [23]. The independent contribution of risk factors like BMI and sex was tested in the development of diabetes using univariate logistic regression. Univariate logistic regression was done for each glucose and insulin concentration independently to predict the development of type-2 diabetes. Based on the Wald test statistic obtained from univariate logistic regression, we identified insulin and glucose concentrations that can independently predict the development of diabetes. Accordingly, two different multivariate logistic regression analyses were then done, one for insulin and another one for glucose concentrations. Multivariate regression analysis for insulin concentration included only those insulin concentrations as predictors that individually predicted ($p < 0.05$) the development of diabetes in univariate logistic regression. The significance of the predictor variable in the model was assessed using the Wald test at 5% level of significance. The multivariate logistic regression for glucose concentrations was done in the same way to identify the glucose concentration that predicted the development of diabetes. Insulin and glucose concentrations that predicted the development of diabetes in multivariate logistic regression along with other predictors like sex and body mass index (BMI) were then used to develop a model for prediction of outcome (diabetes or normal). The predictors that did not predict ($p > 0.05$) the development of diabetes according to Wald statistics were removed from the model and a new model was proposed in which interaction term was also included besides individual predictors. The addition or deletion of interaction term in the model was based on obtaining the lower value of Akaike information criterion (AIC) [24].

To investigate the predictive power of our model we used a classification table and receiver operating characteristics (ROC) curve [25]. The classification table cross-classifies the binary response with prediction of whether the subject develops diabetes or not for some cut-off probability value. Specifically, each cut-off probability value is associated with a particular value of specificity and sensitivity. A decision regarding acceptable levels of sensitivity and specificity involves weighting the consequences of analysis [26]. Sensitivity defined as the proportion of subjects who developed diabetes who were correctly predicted to do so (true-positive test) was calculated as

$$\text{sensitivity} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}} \times 100 \quad (1)$$

Specificity was defined as the proportion of subjects who did not develop diabetes who were correctly predicted not to develop it (true-negative test) calculated as

$$\text{specificity} = \frac{\text{true negatives}}{\text{true negatives} + \text{false positives}} \times 100 \quad (2)$$

We also calculated positive predictive value (PPV) defined as the percentage of subjects correctly predicted to develop diabetes and was calculated as

$$\text{positive predictive value} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}} \times 100 \quad (3)$$

Negative predictive value (NPV) was defined as percentage of subjects correctly predicted not to develop diabetes and calculated as

$$\text{negative predictive value} = \frac{\text{true negatives}}{\text{true negatives} + \text{false negatives}} \times 100 \quad (4)$$

Receiver operating characteristics (ROC) curve is a plot of sensitivity on y-axis and (1 – specificity) on x-axis. Each point on the ROC curve will represent particular sensitivity and specificity value corresponding to a unique cut-off probability value. The area under the ROC curve quantifies how well the model correctly distinguishes a subject with diabetes from a subject without diabetes; the larger the area under the curve or closer the value is to 1, the better the performance of the model. The area under ROC curve ranges from 0 to 1. All data analyses were performed using S-Plus Version 7.0 (Insightful Corporation, Seattle, WA).

1.4. Validation of model

The accuracy of a model is reflected by its ability to predict outcome from other samples of the target population. So, an appropriate procedure is to test the accuracy with new (future) data. However, this method is expensive and delays the final establishment of the model. Also, testing for accuracy with new data is not practical in the present case. The alternative testing is the use of a bootstrap procedure because it reduces the classification bias as compared to jackknife and cross-validation [27]. To validate the model, we used 250 bootstrap repetitions, where sampling was done with replacement. A bootstrap sample “training sample” is drawn from the original data and the model is fit to this sample and the area under ROC curve is used as a measure of predictive accuracy (θ^*). Another bootstrap sample is drawn “test sample” and the area under ROC curve is calculated similar to the calculation for training sample. The process of drawing samples is repeated 250 times and each time the area under ROC curve is calculated. After 250 repetitions, the area under ROC curve for test sample is averaged (θ^*). This produces an average area under ROC curve for training sample (θ^*) and also an averaged area under ROC curve value for test sample (θ^*). The difference in the area under ROC curve between test and training sample ($\theta^* - \theta^*$) gives the ‘bias’. The lower the value of this bias, the better the predictions will be from our model.

We also compared the results from our model to the results that are obtained by using current knowledge about family history of diabetes (both parents are type-2 diabetic) as a risk factor. To estimate the improvement our model is able to make in current knowledge about prediction of type-2 diabetes in subjects with family history of type-2 diabetes, we randomized a group of 191 subjects with family history of type-2 diabetes and assigned diabetic outcome to 25% of the subjects [2] (currently known risk for developing diabetes in subjects with both diabetic parents is 25%). Since we know the true outcome for each subject we can construct a classification table that has number of true-positive, true-negative, false-positive and false-negative subjects. From these values we calculated sensitivity and specificity using Eqs. (1) and (2). Comparison of sensitivity and specificity values from current knowledge to the values from our model provides a metric for improvement in prediction of type-2 diabetes using our

model. We also calculated positive predictive value (PPV) and negative predictive value (NPV) from Eqs. (3) and (4). All validation analyses were performed using S-Plus Version 7.0.

2. Results

The length of follow-up observation was 13 years (12.7 ± 6.4 years) on an average, for all the subjects, regardless of the outcome. Tables 1 and 2 shows the results of univariate logistic regression in which each of the risk factor is a predictor and development of type-2 diabetes is the dependent variable (Figs. 1 and 2). Body mass index (BMI) and sex independently predicted ($p < 0.05$) the development of type-2 diabetes (Table 1). Insulin concentrations at time 0, 20, 30, 40, 50, 60, 90, 120 and 180 min predicted ($p < 0.05$) the development of diabetes while insulin concentrations at time 1, 3, 5 and 10 min did not predict ($p > 0.05$) the development of diabetes in univariate logistic regression according to Wald test statistics (Table 1). Glucose concentrations at time 0, 1, 3, 5, 10, 20, 30, 40, 50, 60, 90, 120 and 180 min predicted the development of diabetes ($p < 0.05$) in univariate logistic regression analysis (Table 2). Model with only baseline characteristics (BMI, sex, age, weight and baseline insulin or glucose concentrations) did not predict the development of diabetes. Multivariate logistic regression analysis involved insulin concentrations at time 0, 20, 30, 40, 50, 60 and 90 min to find insulin concentration that significantly ($p < 0.05$) predicts the development of diabetes. Results from multivariate regression analysis showed that insulin concentration at time 90 min (INS090) best predicted the development of diabetes ($p < 0.05$) according to the Wald test. Glucose concentration at time 90 min (GLU090) was also found to predict the development of diabetes in multivariate logistic regression analysis ($p < 0.05$). Insulin concentration at time 90 min was significantly higher ($p < 0.05$) in diabetic outcome group (46.9 ± 34.3 $\mu\text{U/mL}$) as compared to normal outcome group (21.5 ± 13.7) and glucose concentration at time 90 min was also significantly higher ($p < 0.05$) in diabetic outcome group (84.1 ± 17.9 mg/dL) than in normal outcome group (68.8 ± 13.6). A full model was proposed with development of type-2 diabetes as the dependent variable and BMI, gender, insulin and glucose concentrations at time 90 min as predictor variables (Table 3). Glucose concentration at 90 min and gender were not significant ($p > 0.05$) predictors of development of type-2 diabetes according to the Wald statistics (Table 3) in the full model. This led to a reduced model with only BMI and insulin concentration at 90 min as predictors (Table 4). In this model we also included the interaction term between the predictors. We got slightly lower AIC value after dropping the interaction term between BMI and INS090 (119.5 versus 120). Thus, the final diabetes predictive model was proposed as a combination of BMI and INS090 with the following logistic regression parameters:

$$\text{probability of developing type-2 diabetes } (p) = \frac{1}{1 + e^{-x}} \quad (5)$$

where $x = -5.84 + 0.040$ (insulin concentration at time 90 min in $\mu\text{U/mL}$) $+ 0.116$ (BMI in kg/m^2) (Table 4 and Fig. 3). The calculated cut-off probability (p value) for our model is 0.06, this means that if the probability value calculated from Eq. (5) is ≤ 0.06 , then the subject is predicted to 'not develop' diabetes but if the value of $p > 0.06$, then the subject is predicted to 'develop' diabetes. This cut-off value was decided because it gave us a desired combination of high specificity (99%) and moderate sensitivity (54%).

From the cross-classification table, 29 subjects were correctly classified as diabetic (true-positive), the number of subjects falsely classified as diabetic (false-positive) was 1, while 97 subjects were correctly classified as normal (true-negative), and 25 subjects were incorrectly classified as normal (false-negative). Using Eqs. (1) and (2), sensitivity and specificity were

calculated to be 54% and 99% respectively for these outcomes. From Eqs. (3) and (4), positive predictive value (PPV) and negative predictive value (NPV) was calculated to be 96.7% and 79.5%, respectively. The area under the ROC curve was calculated to be 0.817. When the predictive model was applied to 250 bootstrap samples, the area under ROC curve was calculated as 0.811 from averaged test samples and 0.815 from the training sample. The bias in area under ROC curve was calculated to be $0.815 - 0.811 = 0.004$.

2.1. Comparison to non-pharmacokinetic-based prediction

Based on the current knowledge about development of diabetes in subjects with family history of diabetes (both parents diabetic) as the risk factor, a cross-classification table was made. From this table, the number of subjects classified as true-positive was 14, number of false-positive subjects was 38, number of true-negative subjects was 33 and number of false-negative subjects was 106. From these values, the sensitivity and specificity were calculated to be 11.6% and 46.4%, respectively. The positive predictive value (PPV) and negative predictive value (NPV) was calculated to be 26.9% and 23.7%, respectively.

3. Discussion

Family history of diabetes provides valuable information because it represents the combination of inherited genetic susceptibilities and shared environmental and behavioral factors [28]. The knowledge of family history can be crucial in the prevention, early detection, and treatment of type-2 diabetes [28]. Studies have shown that type-2 diabetes can be prevented or delayed by adopting simple, healthy lifestyle changes, such as a healthy diet and exercise [29–31]. Obesity is defined as BMI ≥ 30 kg/m² but in this analysis, BMI was included as a continuous variable because the predictive value of a continuous variable is higher than a categorical variable. We used insulin and glucose concentrations obtained from an intravenous glucose tolerance test (IVGTT) based on the hypothesis that it may contribute to a better assessment of the risk for developing type-2 diabetes, which is proven in this work. Since glucose and insulin concentrations obtained from the IVGTT test in this study describe the subjects long before the conventional diagnosis of type-2 diabetes, association of these glucose and insulin concentrations with other risk factors provides a new opportunity to predict development of type-2 diabetes. In previous reports, in non-diabetic relatives of subjects with type-2 diabetes, there has been controversy about importance of insulin secretion in predicting development of diabetes [32–35]. Our results show that increased insulin levels is a predictor ($p < 0.05$) for the development of diabetes in subjects with family history of diabetes (both parents type-2 diabetic). Our finding that increased insulin levels at 90 min is a good predictor ($p < 0.05$) for development of type-2 diabetes is an interesting result because insulin concentration at 90 min is a part of the second phase of insulin secretion (from 20 to 120 min). It has been proposed that it is a defect in insulin secretion in the first phase and not the second phase that is predictive of development of type-2 diabetes [36]. Also, predictive factors related to first phase insulin secretion such as total insulin AUC after 5, 10 and 20 min of glucose bolus (including AUC over baseline at all three time points) did not predict the development of type-2 diabetes. Increase in BMI predicted the development of type-2 diabetes ($p < 0.05$). This is expected since obesity increases the risk of developing diabetes. Gender did not play any significant role in predicting the development of type-2 diabetes ($p > 0.05$).

Our final model states that in subjects with family history of diabetes, both higher BMI and increased plasma insulin levels at 90 min are predictors of diabetes. In validation testing, our model maintained good predicting consistency as evidenced by the value of the area under ROC curve and lower bias in area under the ROC curve calculated by bootstrap sampling.

The logistic regression model provided moderate sensitivity (54%) but very high specificity (99%). For our predictive model, the proportion of subjects predicted to develop diabetes who

actually develop it (sensitivity) is important because we want to correctly identify subjects that will develop diabetes. The proportion of subjects predicted not to develop diabetes that actually do not develop it (specificity), was also considered important. The rationale for not ignoring the need for high specificity is that once subjects know that they are less likely to develop diabetes, their life style might become worse with respect to their diet, exercise and other manageable risk factors, and this may put them at higher risk for developing diabetes without this knowledge. Thus, our calculation of the cut-off (classification) probability of 0.06 was based on high specificity (99%) combined with a moderate sensitivity (54%).

Positive predictive value (PPV) of 96.7% means that among subjects with diabetic outcome, 96.7% of subjects were correctly identified as diabetics and negative predictive value (NPV) of 79.5% means that among subjects with normal (non-diabetic) outcome, 79.5% of subjects were correctly identified as non-diabetics. Our logistic regression model is simple to use to predict the development of diabetes based on the probability value (p) and the x value obtained from the model. If $p > 0.06$ (or $x > -2.8$), then the subject is predicted to develop diabetes and if $p \leq 0.06$ (or $x \leq -2.8$), then the subject is not predicted to develop diabetes.

Our model was able to identify 4.5 times more proportion (54% versus 11.6%) of subjects predicted to develop diabetes (true-positive test) and more than two times the proportion of subjects (99% versus 46.4%) predicted not to develop diabetes (true-negative test) than using current knowledge [2] about development of diabetes in subjects with family history of diabetes (both parents diabetic). This substantial improvement in predictability is remarkable since it is based on a single blood sample. No model in the diabetes literature addresses the individual risk assessment in subjects with family history of type-2 diabetes using pharmacokinetic testing. This study is unique in terms of the predictive model that address the individualized risk in subjects with family history of type-2 diabetes based on their BMI and insulin concentration at 90 min. The model is particularly attractive because it only requires a single blood sample at 90 min to determine the plasma insulin concentration. However, this sample comes from IVGTT test that is less practical than oral glucose tolerance test (OGTT). We would like to encourage future OGTT studies to explore our novel model to make individualized predictions development of type-2 diabetes. We believe that the individualized and more accurate risk assessment possible with our model may provide increased incentives to embark on a modified lifestyle, e.g. proper diet and exercise to reduce the risk of developing type-2 diabetes and should lead to a closer attention to the detection of the disease to ensure early medical intervention.

Acknowledgments

We greatly appreciate the kind support of this study by Dr. James Warram, Professor Emeritus of Joslin Diabetes Center, Boston, MA, for making data available.

References

1. Gregg EW, Cadwell BL, Cheng YJ. Trends in the prevalence and ratio of diagnosed to undiagnosed diabetes according to obesity levels in the US. *Diab Care* 2004;27:2806–2812.
2. Annis AM, Caulder MS, Coom ML. Family history, diabetes, and other demographic risk factors among participants of the National Health and Nutrition Survey 1999–2002. *Preventing Chronic Disease*. April;2005
3. Meigs JB, Cupples LA, Wilson PW. Parental transmission of type 2 diabetes: the Framingham Offspring Study. *Diabetes* 2000;49:2201–2207. [PubMed: 11118026]
4. Millar WJ, Young TK. Tracking diabetes: prevalence, incidence and risk factors. *Health Rep* 2003;14:35–47. [PubMed: 12816014]
5. Molyneaux L, Constantino M, Yue D. Strong family history predicts a younger age of onset for subjects diagnosed with type 2 diabetes. *Diab Obes Metab* 2004;187–194.

6. Harrison TA, Hindorff LA, Kim H, Wines RC, Bowen DL. Family history of diabetes as a potential public health tool. *Am J Prev Med* 2003;24:152–159. [PubMed: 12568821]
7. Harwell TS, Dettori N, Flook BN, Priest L. Preventing type 2 diabetes: perceptions about risk and prevention in a population-based sample of adults > or =45 years of age. *Diab Care* 2001;24:2007–2008.
8. Pierce M, Harding D, Ridout D, Keen H, Bradley C. Risk and prevention of type II diabetes: offspring's views. *Br J Gen Pract* 2001;51:194–199. [PubMed: 11255900]
9. Pierce M, Hayworth J, Warburton F, Keen H, Bradley C. Diabetes mellitus in the family: perceptions of offspring's risk. *Diab Med* 1999;16:431–436.
10. Farmer AJ, Levy JC, Turner RC. Knowledge of risk of developing diabetes mellitus among siblings of type 2 diabetic patients. *Diab Med* 1999;16:233–237.
11. Kim J, Choi S, Kim CJ, Oh Y. Perception of risk of developing diabetes in offspring of type 2 diabetic patients. *Korean J Intern Med* 2002;17:14–18. [PubMed: 12014207]
12. Vernon SW. Risk perception and risk communication for cancer screening behaviors: a review. *J Natl Cancer Inst Monogr* 1999;101–119. [PubMed: 10854465]
13. Aiken LS, Fenaughty AM, Luckett TL. Perceived determinants of risk for breast cancer and the relations among objective risk, perceived risk, and screening behavior over time. *Womens Health* 1995;1:27–50. [PubMed: 9373372]
14. Lebovitz HE. Insulin secretagogues: old and new. *Diab Rev* 1999;7:139–153.
15. Porte D Jr, Kahn SE. The key role of islet dysfunction in type II diabetes mellitus. *Clin Invest Med* 1995;18:247–254. [PubMed: 8549009]
16. Warram JH, Martin BC, Krolewski AS, Soeldner JS, Kahn CR. Slow glucose removal rate and hyperinsulinemia precede the development of type II diabetes in the offspring of diabetic parents. *Ann Intern Med* 1990;113:909–915. [PubMed: 2240915]
17. Warram JH, Sigal RJ, Martin BC, Krolewski AS, Soeldner JS. Natural history of impaired glucose tolerance: follow-up at Joslin Clinic. *Diab Med* 1996;13(9 Suppl. 6):S40–S45.
18. Hoffman WS. A rapid photoelectric method for the determination of glucose in blood and urine. *J Biol Chem* 1967;120:51–155.
19. Kadish AH, Little RH, Sternberg JC. A new and rapid method for the determination of glucose by measurement of rate of oxygen consumption. *Clin Chem* 1968;14:116–119.
20. Soeldner JS, Slone D. Critical variables in the radioimmunoassay of serum insulin using the double antibody technique. *Diabetes* 1965;14:771–779. [PubMed: 5850216]
21. Saad MF, Knowler WC, Pettitt DJ, Nelson RG, Mott DM, Bennett PH. The natural history of impaired glucose tolerance in the Pima Indians. *N Engl J Med* 1988;319:1500–1506. [PubMed: 3054559]
22. Edelstein SL, Knowler WC, Hamman RF. Predictors of progression from impaired glucose tolerance to NIDDM: an analysis of six prospective studies. *Diabetes* 1997;46:701–710. [PubMed: 9075814]
23. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. National Diabetes Data Group. *Diabetes* 1979;28:1039–1057. [PubMed: 510803]
24. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr* 1974;19:716–723.
25. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29–36. [PubMed: 7063747]
26. Tabaei BP, Herman WH. A multivariate logistic regression equation to screen for diabetes: development and validation. *Diab Care* 2002;25:1999–2003.
27. Bradley E. The jackknife, the bootstrap, and other resampling plans. *Soc Ind Appl Math* 1982;38:92.
28. Yoon PW, Scheuner MT, Khoury MJ. Research priorities for evaluating family history in the prevention of common chronic diseases. *Am J Prev Med* 2003;24:128–135. [PubMed: 12568818]
29. Tuomilehto J, Lindstrom J, Uusitupa M. Finnish Diabetes Prevention Group. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N Engl J Med* 2001;344:1343–1350. [PubMed: 11333990]
30. Knowler WC, Barrett-Connor E, Fowler SE, Nathan DM. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002;346:393–403. [PubMed: 11832527]

31. Wing RR, Venditti E, Jakicic JM, Lang W. Lifestyle intervention in overweight individuals with a family history of diabetes. *Diab Care* 1998;21:350–359.
32. Weir GC. Which comes first in non-insulin-dependent diabetes mellitus: insulin resistance or beta-cell failure? Both come first. *JAMA* 1995;273:1878–1879. [PubMed: 7776508]
33. Gulli G, Ferranninni E, Stern M, DeFronzo RA. The metabolic profile of NIDDM is fully established in glucose-tolerant offspring of two Mexican-American NIDDM parents. *Diabetes* 1992;41:1575–1586. [PubMed: 1446799]
34. Pimenta W, Korytkowski M, Gerich J. Pancreatic beta-cell dysfunction as the primary genetic lesion in NIDDM. Evidence from studies in normal glucose-tolerant individuals with a first-degree NIDDM relative. *JAMA* 1995;273:1855–1861. [PubMed: 7776502]
35. O'Rahilly SP. Beta-cell dysfunction, rather than insulin insensitivity, is the primary defect in familial type 2 diabetes. *Lancet* 1986;2:360–364. [PubMed: 2874367]
36. Porte D Jr. Banting lecture 1990. Beta-cells in type II diabetes mellitus. *Diabetes* 1991;40:166–180. [PubMed: 1991568]

Fig. 1.

Representative logistic regression analysis plots showing probability of developing diabetes with insulin concentration at 90 min (left plot) and glucose concentration at 90 min (right plot) as the predictor variables. The subjects have only binary outcomes, so original data (represented by circles) is only plotted at those two probability values (0 and 1).

Fig. 2.

Logistic regression analysis showing probability of developing diabetes with BMI as the predictor variable. The subjects have only binary outcomes, so original data (represented by circles) is only plotted at those two probability values (0 and 1).

Fig. 3.

Plot of probability of diabetes with a function of BMI (kg/m^2) and insulin concentration at 90 min ($\mu\text{U/mL}$) as the predictors. X on x -axis is obtained from the final logistic regression model by the following equation: $X = -5.84 + 0.116\text{BMI} + 0.0402\text{INS090}$. For a particular value of X , we can estimate the probability of developing diabetes directly from the plot. The horizontal line represents the probability cut-off value. This cut-off value represents a combination of specificity (99%) and sensitivity (54%). Individual subjects are represented by circles with a probability of developing diabetes as either 0 or 1.

Table 1Univariate logistic regression analysis for development of diabetes^a

Covariate	Coefficient (β)	S.E.	p	Odds ratio ^b	Confidence interval for odds
BMI	0.188	0.0435	<0.001	1.21	1.11–1.32
INS000 ^c	0.0560	0.0153	<0.001	1.06	1.02–1.09
INS001	−0.00109	0.00252	0.67*	1.00	0.99–1.00
INS003	0.000105	0.00240	0.97*	1.00	1.00–1.00
INS005	0.00199	0.00234	0.40*	1.00	1.00–1.00
INS010	0.00440	0.00244	0.07*	1.00	0.99–1.01
INS020	0.00857	0.00310	0.01	1.01	0.99–1.03
INS030	0.0127	0.00360	0.001	1.01	1.00–1.02
INS040	0.0179	0.00435	<0.001	1.02	1.01–1.03
INS050	0.0217	0.00485	<0.001	1.02	1.01–1.03
INS060	0.0280	0.00597	<0.001	1.03	1.02–1.04
INS090	0.0560	0.0106	<0.001	1.06	1.04–1.08
INS120	0.0705	0.0142	<0.001	1.07	1.04–1.10
INS180	0.0667	0.0166	<0.001	1.07	1.04–1.10
SEX ^d	0.740	0.335	0.03	2.10	1.07–4.10

^aModel is defined as (Outcome~Covariate). Each of the covariate is individually modeled as predictor and outcome (diabetes or no diabetes) is the dependent variable.^bFor each covariate, e^{β} is an expression for calculation of odds ratio and odds ratio is defined as the odds of developing diabetes for every unit increase in that covariate.^cINS is insulin concentration ($\mu\text{U/mL}$) and three digits denote time in minutes following the glucose administration.^dSEX is categorical variable: 1 = male, 0 = female.* Not significant ($p > 0.05$).

Table 2

Univariate logistic regression analysis for all glucose concentrations as individual predictors and development of diabetes as dependent variable^a

Covariate	Coefficient	S.E.	p	Odds ratio	Confidence interval for odds
GLU000 ^b	0.0689	0.0212	0.001	1.07	1.63–1.77
GLU001	0.00512	0.00246	0.04	1.00	1.00–1.00
GLU003	0.0167	0.00531	0.002	1.02	1.00–1.03
GLU005	0.0203	0.00588	0.001	1.02	1.01–1.03
GLU010	0.0228	0.00712	0.001	1.02	1.01–1.03
GLU020	0.0370	0.00770	<0.001	1.04	1.02–1.06
GLU030	0.0292	0.00635	<0.001	1.03	1.02–1.04
GLU040	0.0352	0.00700	<0.001	1.04	1.03–1.05
GLU050	0.0324	0.00670	<0.001	1.03	1.02–1.04
GLU060	0.0336	0.00700	<0.001	1.03	1.02–1.04
GLU090	0.0613	0.0118	<0.001	1.06	1.04–1.09
GLU120	0.0522	0.0177	0.003	1.05	1.01–1.08
GLU180	−0.0587	0.0230	0.01	0.94	0.90–0.98

^aModel is defined as (Outcome~Covariate). Each of the covariate is individually modeled as predictor and outcome (diabetes or no diabetes) is the dependent variable.

^bGLU denotes glucose concentration (mg/dL) and three digits are time in minutes.

Table 3Multiple logistic regression analysis with 'full model' for development of diabetes^a

Covariate	Coefficient	S.E.	p	Odds ratio	Confidence interval for odds
Intercept	-8.06	1.87	<0.001	-	-
SEX	-0.181	0.513	0.72*	0.83	0.30-2.33
BMI	0.116	0.0512	0.02	1.12	1.01-1.24
INS090	0.0346	0.0133	0.01	1.04	1.01-1.07
GLU090	0.0345	0.0185	0.06*	1.04	1.00-1.08

^a Full model is defined as (Outcome~SEX + BMI + INS090 + GLU090).* Not significant ($p > 0.05$).

Table 4

Multiple logistic regression using final model^a

Covariate	Coefficient	S.E.	p	Odds ratio	Confidence interval for odds
Intercept	-5.84	1.32	<0.001	-	-
BMI	0.116	0.0494	0.02	1.13	1.02-1.25
INS090	0.0402	0.0128	0.002	1.49*	1.45-1.53*

^aFinal model is defined as (Outcome~BMI + INS090).

* Estimated odds ratio and 95% confidence intervals for 10-unit increase; INS090 is insulin concentration at time 90 min (in μU/mL).