Can J Diabetes xxx (2016) 1–5

Contents lists available at ScienceDirect

Canadian Journal of Diabetes

journal homepage: www.canadianjournalofdiabetes.com

A Simple Score Model to Assess Prediabetes Risk Status Based on the Medical Examination Data

Peng Ouyang PhD Candidate [a], Xitong Guo PhD [a], Yiting Shen MSc [b], Naiji Lu PhD [a,c,*], Chenghua Ma MSc [a]

[a] School of Management, Harbin Institute of Technology, Harbin, China
[b] School of Management, University of Bath, Bath, United Kingdom
[c] Department of Biostatistics, University of Rochester, Rochester, New York, United States

ARTICLE INFO

Article history:
Received 17 November 2015
Received in revised form
23 February 2016
Accepted 23 February 2016

Keywords:
medical examination data
prediabetes
risk categorization
risk score model

Mots clés :
données des examens médicaux
prédiabète
catégorisation des risques
modèle de cotation du risque

ABSTRACT

Objectives: We aimed to build a risk score model to screen out the patients at high-risk status so as to prevent or delay the conversion of prediabetes to diabetes.
Methods: The population were divided into 2 groups: 1 was an exploratory population, and the other was a validation population. All the data were extracted from the electronic medical examination datasets in the School Hospital of Harbin Institute of Technology, Harbin, China. A binary logistic regression model was used to screen out the risk factors, and the associated risk factors were categorized into 3 levels to create the prediabetes score model. We divided the total score into 4 risk categories: low, middle, high and extremely high risk. We also tested the performance of our prediabetes risk score model.
Results: Age, body mass indexes, histories of hypertension, family histories of diabetes, diastolic blood pressure levels and triglyceride levels were screened out as independent risk factors in order to build the risk score model. The area under the curve (AUC) of the prediabetes risk score model was 0.748 (95% CI, 0.720 to 0.777), and the AUC for the validation population reached 0.713 (95% CI, 0.686 to 0.740). Low, middle, high and extremely high risk statuses for prediabetes were associated with a total score of 0 to 3, 4 to 6, 7 to 10 and 11 to 12.
Conclusions: Our prediabetes score model can be used easily and understood by doctors and other related users to assess prediabetes risk status. The intervention program, designed based on our prediabetes score model, is likely to prevent or delay the conversion of prediabetes to diabetes.

RÉSUMÉ

Objectifs : Nous avions pour objectif d'élaborer un modèle de cotation du risque pour déterminer les patients exposés à un risque élevé afin de prévenir ou de retarder la conversion du prédiabète en diabète.
Méthodes : Nous avons divisé la population en 2 groupes : le 1er comptait une population de type exploratoire et le 2e comptait une population de validation. Toutes les données ont été extraites des ensembles de données électroniques des examens médicaux du School Hospital of Harbin Institute of Technology, de Harbin, en Chine. Un modèle de régression logistique binaire a été utilisé pour déterminer les facteurs de risques, et les facteurs de risque associés ont été répartis en 3 catégories pour créer le modèle de cotation de prédiabète. Nous avons divisé le score total en 4 catégories de risque : faible, moyen, élevé et extrêmement élevé. Nous avons également vérifié la performance de notre modèle de cotation du risque de prédiabète.
Résultats : L'âge, les indices de masse corporelle, les antécédents d'hypertension, les antécédents familiaux de diabète, les valeurs de pression artérielle diastolique et les concentrations de triglycérides ont été déterminés comme facteurs de risque indépendants afin d'élaborer le modèle de cotation du risque. La surface sous la courbe (SSC) du modèle de cotation du risque de prédiabète était de 0,748 (IC à 95 %, 0,720 à 0,777), et la SSC de la population de validation atteignait 0,713 (IC à 95 %, 0,686 à 0,740). Les niveaux de risque faible, moyen, élevé et extrêmement élevé de prédiabète ont été associés à un score total de 0 à 3, de 4 à 6, de 7 à 10 et de 11 à 12.

* Address for correspondence: Naiji Lu, PhD, Harbin Institute of Technology, 92 West Dazhi Street, Nan Gang District, Harbin, China.
E-mail address: nlu@hit.edu.cn

1499-2671 © 2016 Canadian Diabetes Association.
http://dx.doi.org/10.1016/j.jcjd.2016.02.013

Conclusions : Notre modèle de cotation de prédiabète peut être utilisé facilement et compris par les médecins et les autres utilisateurs pour évaluer le niveau de risque de prédiabète. Le programme d'intervention conçu selon notre modèle de cotation de prédiabète est susceptible de prévenir ou de retarder la conversion du prédiabète en diabète.

Introduction

Diabetes, a prevalent chronic disease, and its related complications are becoming an urgent worldwide health issue and a serious public health problem in many countries, including Canada, the United States, China, Japan and the United Kingdom. The incidence of diabetes was 382 million in 2013, and it is projected that this number will move up to 592 million in 2035 if there is no appropriate prevention (1). In China, 9.7% of adults had total diabetes, and 15.5% of adults had prediabetes, which means 92.4 million adults have diabetes, and 148.2 million adults have prediabetes (2). Diabetes, often accompanied by hypertension (3), is associated with cardiovascular complications, i.e. heart failure (4), accelerated atherosclerosis and cardiac dysfunction (5). Diabetes is a leading cause of mortality and mobidity. It is expected that total death rates resulting from diabetes will increase more than 50% in the next 10 years (6). This indicates that diabetes will be a huge burden for the healthcare system, given the scarcity of healthcare resources and the potential probability that prediabetes will progress to diabetes.

Prevention is considered to be the best intervention because diabetes cannot be cured. Some studies of lifestyle intervention programs in the United States (7), China (8) and Europe (9) have shown that it is possible to reduce the incidence of type 2 diabetes by about 50% in at-risk individuals. So screening individuals at high risk for developing diabetes is of significant importance. Prediabetes, not a disease but a high-risk state of conversion to diabetes, is defined as a situation in which blood glucose concentrations are higher than normal but lower than the diabetes threshold (10,11). Some evidence has shown that lifestyle modifications, as the conerstone of diabetes prevention, could lead to a 40% to 70% relative-risk reduction (12). Compared with interventions after the onset of diabetes, the related intervention measures started during the prediabetes stage are more efficient and sensible because they can prevent or delay the conversion from prediabetes to diabetes (13,14).

The risk score model is a practical and efficient tool to assess individuals' diabetes development status and screen out the target population at high risk. Currently, there exist many diabetes risk score models (15–18) that could be used to optimize the estimation of diabetes risk and help physicians to evaluate patients' risk status and make related intervention plans. However, the risk model for prediabetes is rather rare. To the best of our knowledge, we have only 1 prediabetes risk score model (19) and 1 prediabetes risk calculator tool (20), and the only risk score model targets children.

The aim of our research was to bulid a prediabetes risk score model based on the data in the medical examination reports. We anticipated that our prediabetes risk score model could help physicians to assess the progress of the conversion of prediabetes to diabetes and make related intervention plans to help patients prevent or delay the onset of diabetes.

Methods

Study design and population

This study was performed on the basis of medical examination data that were sourced from annual check-ups in the School Hospital of Harbin Institute Technology in 2014. We extracted sex, age, body mass index (BMI), history of hypertension, history of coronary heart disease, family history of hypertension, family history of diabetes, family history of cerebrovascular disease, history of cerebrovascular disease, smoking or not, drinking or not, systolic blood pressure, diastolic blood pressure, total cholesterol, triglyceride, high-density lipoprotein cholesterol (HDL-C) and low-density lipoprotein cholesterol (LDL-C) as potential risk factors. We excluded persons if any data were missing for the above risk factors.

The goal of our study was to build a prediabetes score model. For the purpose of our study, the criteria of prediabetes was defined as individuals with fasting plasma glucose (FPG) levels of 5.6 to 6.9 mmol/L, according to Canadian Diabetes Association Clinical Practice Guidelines Expert Committee criteria (21). We excluded patients whose FPG levels were more than 7.0 mmol/L because they could be diagnosed as having diabetes, according to the definition of diabetes; however, patients with diabetes were not our targeted population. There are 3099 patients in our population. We randomly divided the population into 2 groups: 1 was an exploratory population including 1360 patients, and the other was validation population including 1739 patients. The exploratory population was used to build the risk score model; the validation population was a test group to examine the performance of our risk score model. To test the performance of the risk score model, we included more patients in the validation group.

The research was approved by the ethics committee of the School Hospital of Harbin Institute of Technology. For privacy, all of the names and the medical exmination document numbers were deleted by the School Hospital of Harbin Institute of Technology.

Statistical analysis

We performed our statistical analysis work in the R program (R Foundation for Statistical Computing, Vienna, Austria) (22). For the demographic characteristics and clinical biochemistry expression, we used mean and standard deviations. We set our dependent dummy variable according to the following rules: if FPG levels were in the prediabetes range, the dependent variable was set as 1; if not, it was 0. History of hypertension, history of coronary heart disease, family history of hypertension, family history of diabetes, family history of cerebrovascular disease, history of cerebrovascular disease, smoking or not and drinking or not were treated as dummy variables.

We put the extracted data variables, including sex, age, BMI, history of hypertension, history of coronary heart disease, family history of hypertension, family history of diabetes, family history of cerebrovascular disease, history of cerebrovascular disease, smoking or not, drinking or not, systolic blood pressure, diastolic blood pressure, total cholesterol, triglyceride, HDL-C and LDL-C into a binary logistic regression model through stepwise backward methods. The variables in the logistic regression model were screened out to build the risk score model. For the selected significant variables, we plotted the receiver operated characteristic (ROC) curve to get the cutoff value of each index through the Youden index, which maximizes the vertical distance from the line of equality to the cutoff value point. That is to say, the Youden index is the point on the ROC curve that is the farthest from the line of equality. The selected risk variables were divided into 3 levels in order to formulate the related scale, except for the dummy variables. Then we put the scale variables into a binary logistic regression model to get the coefficients. The score value was assigned based on the