# MAIN PREDICTORS OF DIABETES MELLITUS TYPE 2 (T2DM) IN THE RUSSIAN FEDERATION

M. Shestakova[1], A. Mayorov[1], G. Galstyan[1], A. Kolbin[2,3], M. Arepyeva[3], M. Proskurin[3], O. Karpov[4], A. Kurylev[2]

[1] Endocrinology Research Center , Moscow, Russia; [2]1-st Saint-Petersburg State Medical University, Saint-Petersburg, Russia; [3] Saint-Petersburg State University, Saint-Petersburg, Russia; [4] Sanofi, Moscow, Russia

## BACKGROUND:

The prevalence of T2DM has been significantly increasing both in the Russian Federation (RF) and worldwide. The RF data from a federal registryindicate that there were 4.35 mln people (3.0% of the population) registered including 4 millions (92%), 255 thousands (6%), and 75 thousands (2%) people with T2DM, T1DM, and other DM types, respectively. The results of the large-scale Russian epidemiologic study NATION (26,62 thousands records, 2016 year) demonstrated that T2DM is diagnosed only in 46% of patients, prevalence was calculated as 5,4% in adult population[1]. In total there are at least 8–9 million people with T2DM in the RF, which is a great problem in a long-term perspective since a great number of patients are undiagnosed. Consequently, they do not receive treatment and have a high risk of vascular complications.

## AIM:

Main predictors of T2DM detection which could be used for screening of T2DM based on NATION results.

## METHODOLOGY:

A factor analysis of the pooled anonymized NATION database revealed 19 potentially relevant factors that correlate with T2DM among all predictors. All database factors were used as binary variables and processed using a logistic regression method **(Fig. 1).** Logistic regression was used to compute coefficients for risk factors for diabetes. Different combinations of factors were tested and used to train a logistic regression model in order to identify the optimal set of factors determining the maximum area under the receiver operating characteristic curve (ROC-curve). The sensitivity analysis was used to group continuous variables into binary variables. Grouping thresholds were set so as to maximize the area under the ROC-curve of the established model based on 10 predictors (13 predictors in total). Two thersholds were used for age, BMI, and waist circumference. The optimal model was cross-validated.

## RESULTS:

In total 22 parameters and 4 194 304 their combinations form NATION database were obtained. The following parameters were included into the model on the first stage: age, BMI , abnormal fasting blood glucose test, systolic blood pressure, low annual income, waist-to-hip ratio, alcohol consumption, frequency of fast-food intake, sugar-sweetened beverages, complications (atherosclerosis, cerebrovascular disorders, renal insufficiency etc.) etc. After mathematical modelling 16 384 variations of combinations were obtained. The following variables showed strong correlation with T2DM and were included in an optimal model: age, body mass index, waist circumference, at least one reading of elevated blood sugar (fasting), low physical exercise / activity, high waist-to-hip ratio, arterial hypertension, high systolic blood pressure, use of lipid-lowering drugs, certain comorbidities, T2DM in close relatives. The most influential factors with positive effect were: elevated blood glucose, age, BMI, waist-to-hip ratio and systolic blood pressure. Low physical activity had a pronounced negative effect.

The following thresholds for T2DM detection were set based on logistic regression model with corresponding coefficients: for BMI (BMI 26-30, logistic regression coefficient = 0.59;BMI> 30, coefficient = 1.19), for age (49-51 years for women and 46-58 years for men, coefficient = 0.82; age > 51 years for women and >58 years for men, coefficient = 1.24) and waist-to-hip ratio (0.85 for women and 0.92 for men, coefficient = 1.05) (Fig. 2-4). Family history of diabetes had a strong correlation with diabetes diagnosis. Results for T2DM logistic regression model: sensitivity = 83.9%, specificity = 76.8%, Positive Predictive Value (PPV) = 12.3%, Negative predictive Value (NPV) = 99.2%, area under the ROC-curve 0.88.
Results for the prediabetes logistic regression model: sensitivity = 75.1%, specificity = 71.5%, PPV = 40.3%, NPV = 91.8%, area under the ROC-curve 0.802.

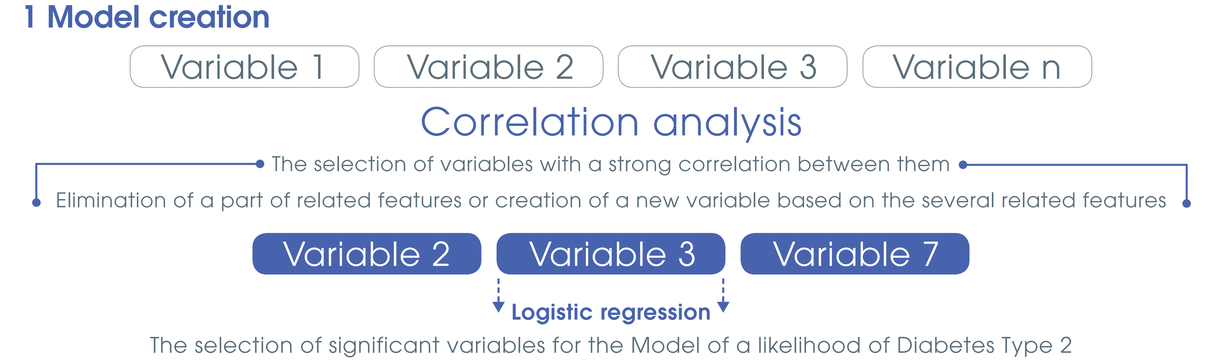
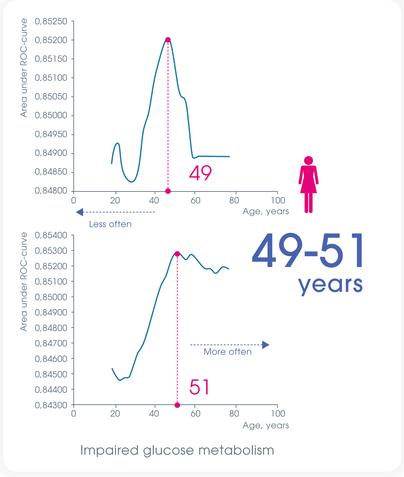



**Fig. 1.** Methodology of analysis.



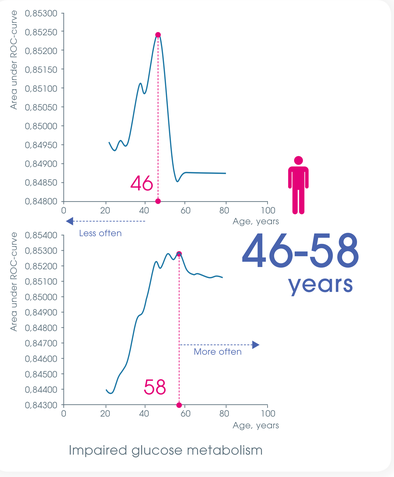




**Fig. 2.** Age thresholds for the the impaired glucose metabolism.

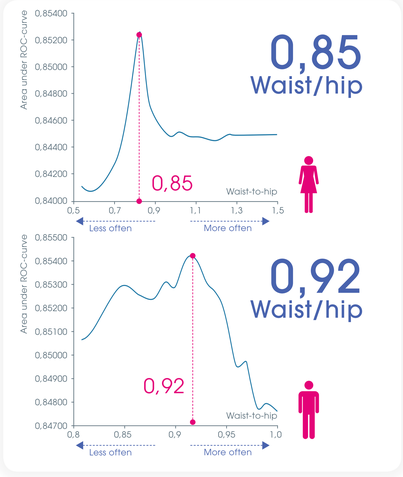



**Fig. 3.** Waist-to-hip ratio thresholds for the impaired glucose metabolism

**The model is the same on all charts, all parameters except the variable are fixed.**

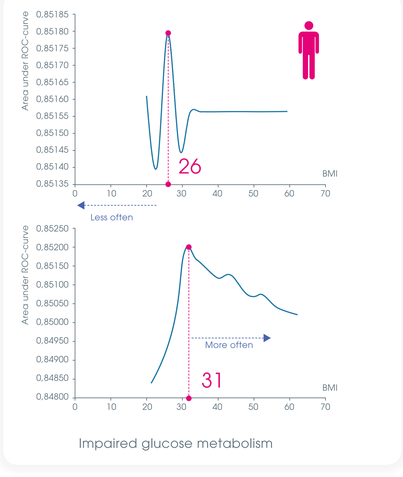




**Fig. 4.** Body mass index (BMI) thresholds for the impaired glucose metabolism

## CONCLUSION:

1. The most important predictors for T2DM in the Russian population are age, BMI, waist-to-hip ratio, family history of Diabetes.
2. The constructed logistic regression model showed sufficient sensitivity and specificity to be used as a basis for a screening tool for impaired glucose metabolism.

### References:

[1] Dedov I, Shestakova M, Benedetti MM, et al. Prevalence of type 2 diabetes mellitus (T2DM) in the adult Russian population (NATION study). Diabetes Res Clin Pract. 2016;115:90-95. doi: 10.1016/j.diabres.2016.02.010

Sanofi
Tverskaya str., 22, Moscow, 125009, Russia
Tel (495) 721 14 00
SARU.LALI.19.09.1612
Conflict of interests: O.Karpov - employee of Sanofi, all others authors have declared about absent of conflict of interests

Sponsored by Sanofi