ADAPTING NEURAL NETWORK MODELS TO PREDICT 10-YEAR CVD DEVELOPMENT BASED ON REGIONAL DATA CALIBRATION

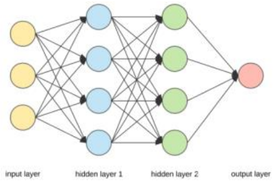
I. Korsakov[1], D. Gavrilov[1], L. Serova[1], A. Gusev[1], R. Novitskiy[1], T. Kuznetsova[2]

(1) K-SkAI, (2) Petrozavodsk State University, Petrozavodsk, Russian Federation

OBJECTIVE The tools for the cardiovascular risk prediction using machine learning methods have proven better prognostic value in comparison with commonly used scores (e.g., Framingham). To improve the accuracy of models the data with regional characteristics: ethnic, nutrition, climatic conditions, living standards, etc, can be considered. These features could significantly affect the development and outcomes of CVD

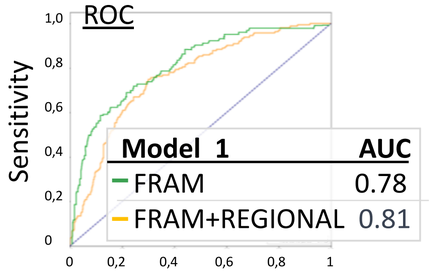
PURPOSE Creation the machine learning models based on publicly available data and validation them on regional medical data

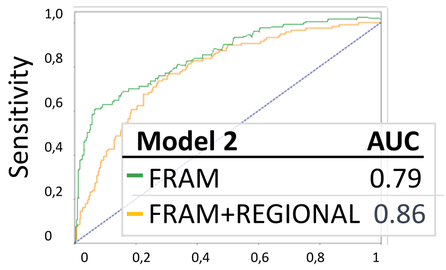
EVALUATION The dataset was divided into Training set (80%) and Validation set (20%). The models were implemented with Keras convolution neural network using 4 layers. For data validation was used a 10 K-fold method



Neural Network

RESULTS We compared the initial model metrics and those obtained after regional data retraining





ESC Congress 2020
The Digital Experience

METHODS 2 models were trained using data from the **FRAMINGHAM** STUDY :

	PATIENTS FROM FRAMINGHAM (n)	INPUT (FEATUREAS)	OUTCOMES
MODEL 1	2 588	age, gender, cholesterol, HDL, smoking, SBP, and BP medications	10-year CVD
MODEL 2	4 363	age, gender, cholesterol, smoking, SBP, BMI, heart rate	10-year death from CVD

REGIONAL dataset collected from 438 patients in North Western Russia was used to **retrain** the obtained models. This data set includes CVD and death from it during a 10-year follow-up

	PATIENTS (n)	AGE (y.o)	MALE (%)	SMOKE (%)	TOT CHOL (mg/dl)	HDLC (mg/dl)	SYSBP (mmHg)	BPMEDS (%)	BMI	HEART RATE (beats per minute)	CVD (%)	CVD DEATH (%)
FRAMINGHAM DataSet	4434	49,9 ± 8,6	56	49	6,16 ± 1,16	1,26± 0,39	132,8± 22,3	55	25,8 ± 4,1	75,8 ± 12,0	23,4	19,5
REGIONAL DataSet	438	47,9 ± 5,7	58,9	37,7	5,16 ± 0.75	1,32± 0.30	128,3± 14,8	21,2	27,5 ± 4,4	57.0 ± 21,7	26,8	21,1

CONCLUSION Using this method of retraining predictive models, we can take into account local characteristics of the population and significantly increase the Accuracy of predicting CVD events

Declaration of interest: I have nothing to declare