



## Non-invasive Predictive AI Coronary Angiography

**TP Abdualimov<sup>1\*</sup> and AG Obrezan<sup>1,2</sup>**

<sup>1</sup>Department of General Practice and Therapy, My Medical Center, Saint-Petersburg, Russian Federation

<sup>2</sup>Department of Hospital Therapy, St. Petersburg State University, Saint-Petersburg, Russian Federation

**\*Corresponding Author:** TP Abdualimov, Department of General Practice and Therapy, My Medical Center, Saint-Petersburg, Russian Federation.

**Received:** January 27, 2023

**Published:** February 16, 2023

© All rights are reserved by **TP Abdualimov and AG Obrezan.**

### Abstract

A novel approach to diagnosing coronary artery disease was proposed. A model for diagnosing coronary heart disease was designed using neural network analysis and allow to reveal transient myocardial ischemia, pathology of the main coronary arteries. The aim of the study was to compare the accuracy of the trained neural network model on the input structured data (sex and age, cholesterol levels, presence of chronic diseases, hereditary factors, lifestyle and etc.) and ECG images with the results of traditional coronary angiography. The proposed diagnostic model was proved to be reliable and highly sensitive for 1500150 cases. The model was compared with the traditional diagnostic methods of transient myocardial ischemia (24-hour Holter monitoring, treadmill test), where the presented diagnostic model was considered to be significantly effective. The accuracy of forecasts was assessed and justified by the cardiologists supervising patients with ACS on a daily basis. The study also presents a new method of sample extrapolation using generative adversarial networks allowing to exceed the volume of observations used in classical meta-analyses.

**Keywords:** Coronary Arteries; Neural Networks; Artificial Intelligence; Coronary Heart Disease; Deep Learning; ECG; Non-invasive Predictive AI Coronary Angiography

### Abbreviations

AV: Atrioventricular; CABG: Coronary Artery Bypass Grafting; CA: Coronary Angiography; ESC: European Society of Cardiology; AI: Artificial Intelligence; ACS: Acute Coronary Syndrome; ACI: Acute Cerebrovascular Insufficiency; CCS: Chronic Coronary Syndrome; ECG: Electrocardiogram; AUC: Area Under Receiver Operating Characteristic Curve

### Introduction

Ischemic heart disease and other pathologies of the cardiovascular system remain the main causes of death in people around the world. According to statistics, 621 people die from cardiovascular diseases in Russia every 100,000 people.

Unfortunately, in the regions this figure is much higher than in the central part of our country.

The “gold standard” for diagnosing these diseases is coronary angiography, an X-ray contrast method of examination that allows you to reliably assess the condition of the coronary arteries that deliver oxygenated blood to the heart, ensuring its smooth operation. However, classical coronary angiography is an invasive technique that has a number of contraindications and, like any surgical procedure, carries certain risks. There are other, non-invasive types of diagnostics (for example, CT angiography, MRI of the heart, 24-hour ECG monitoring, etc.), but they all require expensive equipment and a long stay of the patient in the clinic.

In addition, these conditions often develop asymptotically, so patients go to the doctor already with such acute, life-threatening forms as unstable angina and myocardial infarction. There are a number of other difficulties associated with the detection of heart vessel pathologies: mass self-diagnosis of patients using information from Internet sources, the complexity of clinical diagnosis and verification of the diagnosis by a doctor.

In order to clarify the possibility of the effectiveness of the use of artificial intelligence in assessing the pathology of the coronary bed and to solve the problems described above, we conducted this study.

The promise of AI and machine learning in cardiology is to provide a set of tools to improve the efficiency of the cardiologist.

The introduction of technologies such as polygenome sequencing and streaming of biometric data from mobile devices into clinical practice will soon require cardiologists to interpret and apply information from many disparate areas of biomedicine [1-4].

At the same time, the growing workload in medicine requires doctors and health care systems to become more operationally efficient [5]. Finally, patients are beginning to demand a faster and more personalized approach [6,7]. The amount of data that a specialist has to work with is increasing, more complex interpretation is required, and an increase in the efficiency of physicians is expected [8,9]. The solution is machine learning, which can improve every step of patient care, from research and discovery to diagnosis and therapy selection (Figure 1).

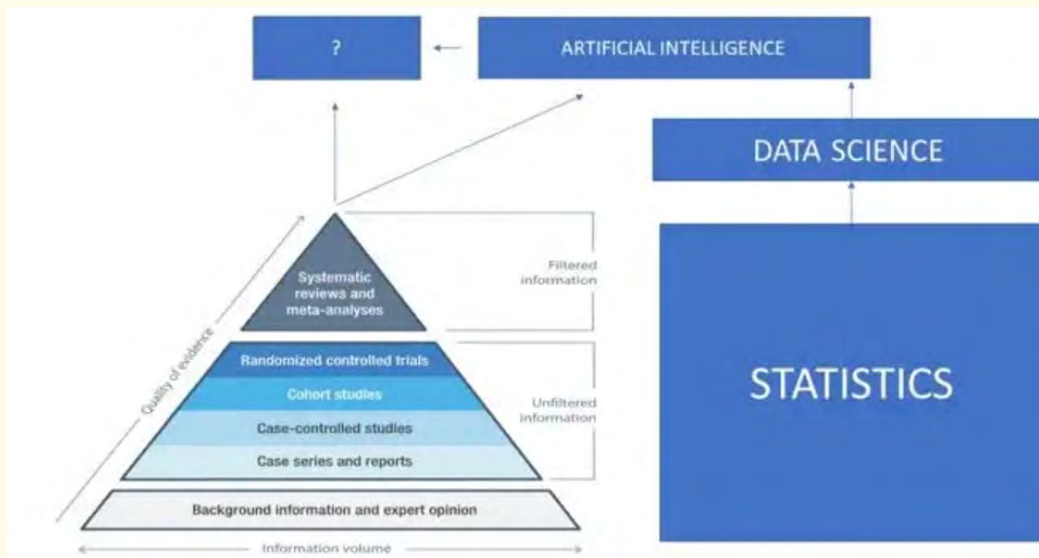


Figure 1: Place of data science in evidence-based medicine.

### Materials and Methods

To train the neural network, a database with ECG images was collected; for this, 100 patients took part in the study, who underwent coronary angiography in a planned and emergency order. Based on the data of 100 patients, a neural network was trained. A supervised learning algorithm was used, in which the outcomes (coronarogram data) were known, and the neural network parameters were adjusted so as to minimize the error.

The indications for coronary angiography were verified according to the recommendations of the ESC. The study was carried out in accordance with Good Clinical Practice and Declaration of Helsinki principles. Inclusion and exclusion criteria were defined.

### Key inclusion criteria

- Signing the informed consent prior to the study, including the statistical processing of medical history data;
- Aged over 18 years;

- Indications (elective or emergency) for coronary catheterization;
- Recorded electrocardiography (25 mm/s) one day before/or less before the performed coronary catheterization.

#### Key exclusion criteria

- ECG identification of arrhythmias as the atrial fibrillation, AV nodal reentrant tachycardia, ventricular tachycardia while recording;
- Previous stenting and/or coronary artery bypass grafting;
- Pronounced disturbances on the recorded ecg;
- Registration of ecg more than 24 hours before coronary catheterization;
- Any surgical or medical state that, according to the researcher, could significantly interfere the work of machine learning algorithm in relation to the accuracy of the results.

The doctor, conducting the study, analyzed the medical record data (complaints, anamnesis, objective, laboratory and instrumental data) and download these results into a machine learning database in a binary format.

At the first stage of data collection for each case, structured parameters were entered into a tabular form, as well as the ECG image in jpeg format into the database. Numerous morphometric, objective, laboratory and instrumental data of the patients were used to train neural networks, such as : age, gender, diagnosed ACS or CCS, ST segment pathology on the ECG, the presence or absence of concomitant pathology (diabetes mellitus, hypertension, obesity, anemia, previous stroke, atherosclerosis, arrhythmias, dyslipidemias), aggravated heredity, bad habits (smoking, alcohol abuse), stress factors, low physical activity, menopause, increased nutritional intake.

The abovementioned factors were filled in a structured binary form (0, 1) in a tabular format. Registration of ECG on a sample, developed for a neural network training, was carried out using one type of apparatus and the record was transmitted to the machine learning operator in jpeg format. Thus, 22 parameters (key features) were used to develop a neural network learning algorithm.

Neural network was trained on the data obtained from the analysis of coronary angiograms. As “targeted” values were taken:

- Performed stenting or recommended CABG based on coronary catheterization,
- Atherosclerosis,
- Left main coronary artery stenosis,
- Left main coronary artery subocclusion,
- Anterior interventricular artery occlusion,
- Anterior interventricular artery subocclusion,
- Anterior interventricular artery stenosis,
- Circumflex artery occlusion,
- Circumflex artery subocclusion,
- Circumflex artery stenosis,
- Right coronary artery occlusion,
- Right coronary artery subocclusion,
- Right coronary artery stenosis.

The degree of the coronary artery stenosis was filled in the table in numerical form as a percentage, then converted to binary form (1 — stenosis more than 50%), the rest of the parameters were filled in binary form according to the presence or absence of lesion. The above “target” values were predicted by the trained machine learning algorithm on three samples.

The algorithm needed to solve the problem of classifying coronary artery lesions, predict the absence or presence of stenosis and their severity. To solve the problem of classifying coronary artery lesions according to the “0;1” system, a neural network was used that takes structured data and an image as input, and a multifactorial classification of coronary arteries was obtained at the output. As software for building the architecture of a neural network, sets of libraries for the Python programming language were used (pandas — for working with tabular data; tensorflow — for designing neural networks and training them).

The input of the neural network simultaneously received ECG images of size (200, 200, 1) and structured tabular data. At the output, the neural network predicted multilevel values of affected coronary values in a probabilistic form. Fully connected, convolutional, batch normalizing (batch normalization layer), “dropout” (exclusion layer) were taken as neural network layers for image processing. For processing structured data, only fully

connected layers are taken. Inside the neural network, a connecting “concatenate layer” was used to generalize the weights of the image and the dataset. After the generalizing layer, there are two

fully connected layers. The output layer consists of 13 neurons for predictions for each parameter (Figure 2).

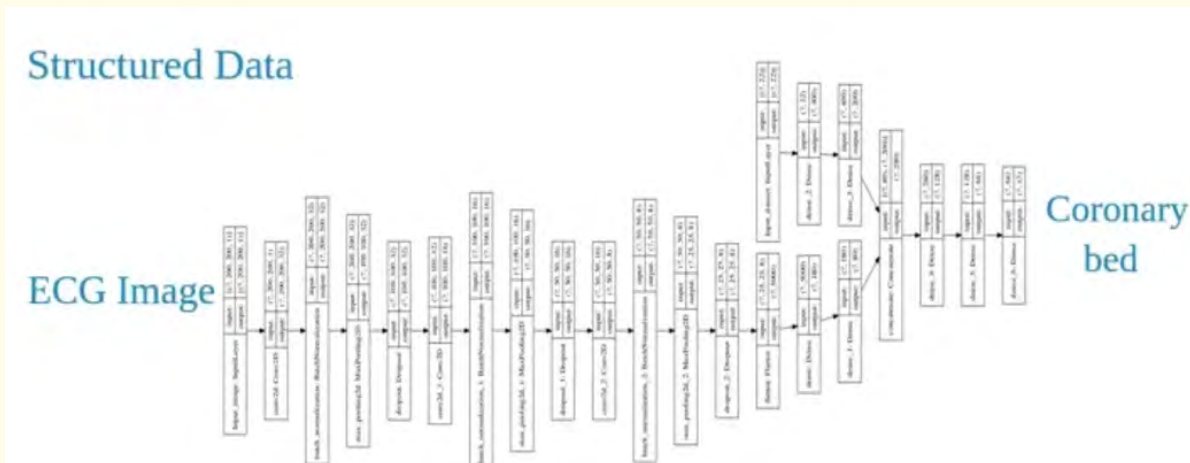


Figure 2: The structure of the neural network.

“Adam” (adaptive learning rate optimization algorithm by calculating the exponential moving average gradient and quadratic gradient) was taken as the optimizer, the loss function is binary crossentropy. Training was performed on 100 “epochs” (one epoch — one forward pass and one reverse pass of all training examples). The size of the “batch” (the number of training examples per iteration) is 8, the size of the validation set is 0.1. The selection of parameters and structure of the neural network was made empirically. AUC was chosen as the starting metric for assessing the quality of the model.

The assessment of accuracy was made on specially selected test samples, the comparison was carried out according to the data of coronary angiograms obtained during the performance of invasive coronary angiography (Figure 3).

Sample 1

20 inpatients with an extremely complex and atypical clinical picture, features of the anatomy of the coronary bed. Example 1: an elderly patient with a typical clinical picture of anginal pain, risk factors — according to coronary angiography without pathology. Example 2: an elderly patient without anginal pain, previously verified atherosclerosis — according to coronary angiography,

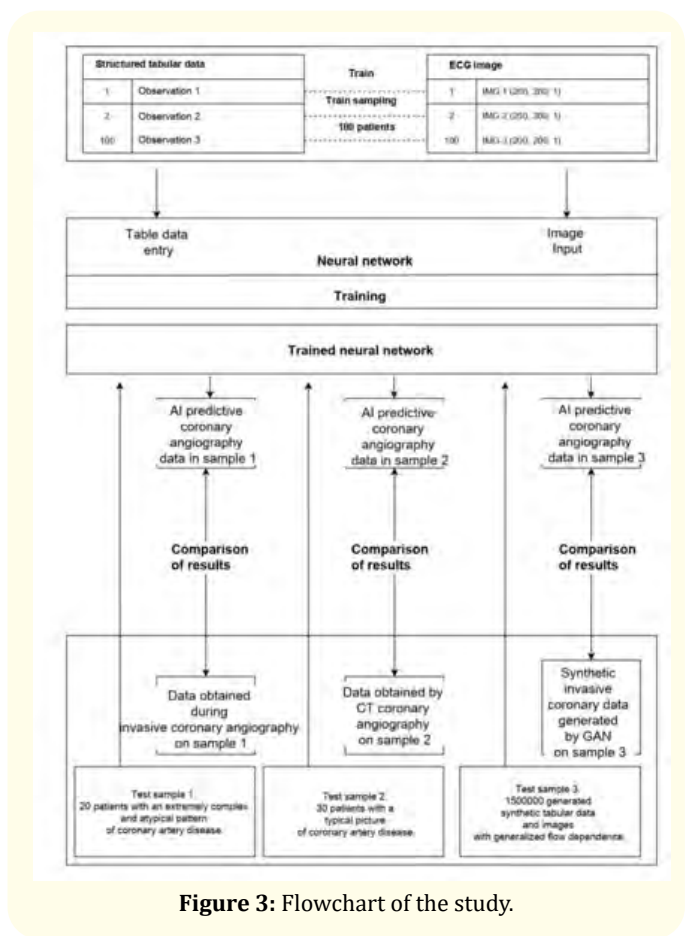


Figure 3: Flowchart of the study.

a multi-vessel lesion of the coronary bed with involvement of the trunk. On the data of 20 patients, 20 tasks were compiled. Cardiologists who supervise daily patients with ACS were asked to predict the presence of myocardial ischemia, damage to the main arteries. Comparison of accuracy with the trained neural network is carried out.

**Sample 2**

30 outpatients with a typical clinical picture or no symptoms of coronary artery disease. The accuracy of the trained neural network was compared with the results of CT angiography. Before performing CT coronary angiography, patients underwent a treadmill test and daily ECG monitoring. The accuracy of detection of transient myocardial ischemia was compared with classical methods.

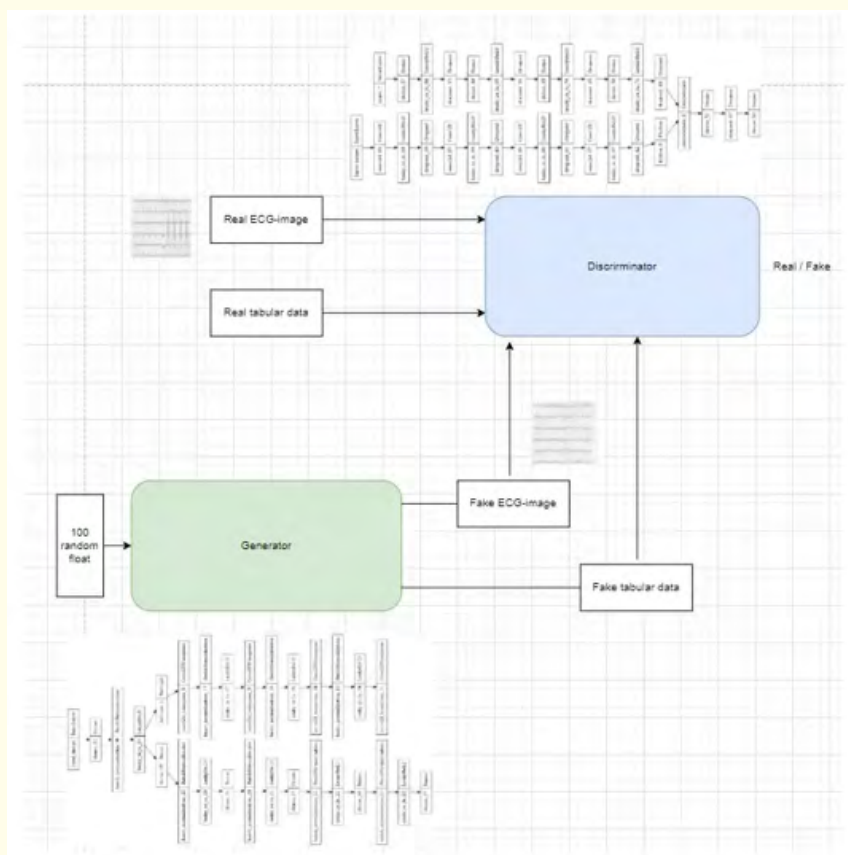
**Sample 3**

The authors were inspired to create this sample by an article by colleagues [10]. In this publication, signs of a new coronavirus infection were detected on radiographs using neural network analysis. The authors, using GAN, generated x-ray images with lesions characteristic of a new coronavirus infection, and on these images,

along with real x-rays, trained a neural network and achieved good results. We were interested in checking how accurate the trained neural network is on an extremely large sample generated by the GAN.

100 random numbers with a normal distribution were fed to the generator input. The output generated an image (200, 200) and structured tabular data of size (1, 35). (one row, 35 columns). There was a generalization layer inside the generator to keep the data streaming between the table row and the image. The input of the discriminator was a generated image of size (200, 200) along with real ECG images (200, 200) and generated tabular data of size (1, 35) along with real tabular data. At the output, the discriminator produced a binary classification corresponding to real data and synthetic ones.

Thus, it was necessary for two neural networks to outperform each other. One neural network tried to generate an image and a table that the discriminator did not distinguish from real ones, tried to look for features characteristic of a real image and table in order to distinguish the generated images and table from real ones (Figure 4).



**Figure 4:** Structure of a generative adversarial neural network.



1500000 ECG images and structured data were obtained (1500000\*35 table). After generation, the authors needed to solve the problem. To what extent the generated data is similar to the real one and whether the flow dependence of features is preserved.

For ECG images, conventional visual analysis was used. ECG images are outwardly almost indistinguishable from real ones (Figure 5).

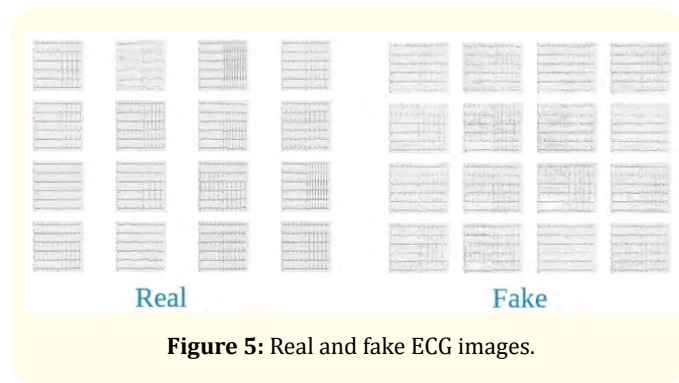


Figure 5: Real and fake ECG images.

A more difficult problem is tabular data, how close are they to real ones?

The distribution of patients by age was analyzed. The distribution of real data is normal, synthetic data are distributed with three peaks, towards the median, minimum and maximum values (Figure 6).

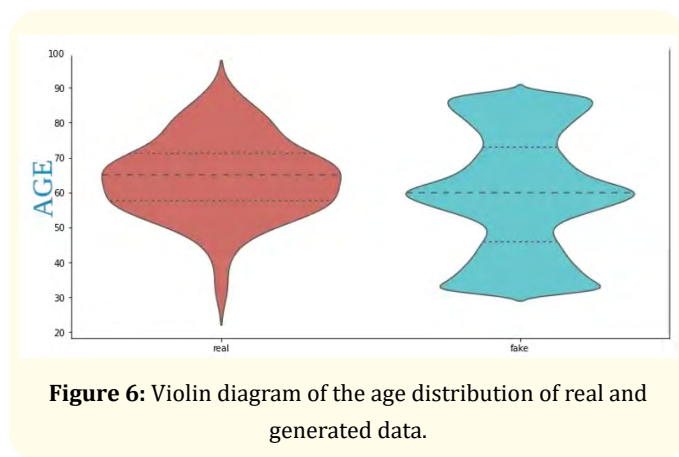


Figure 6: Violin diagram of the age distribution of real and generated data.

A quantitative analysis of the generated features was performed. The distribution is close to the real one, however, the revealed differences are in the quantitative relation of the signs (Figure 7).

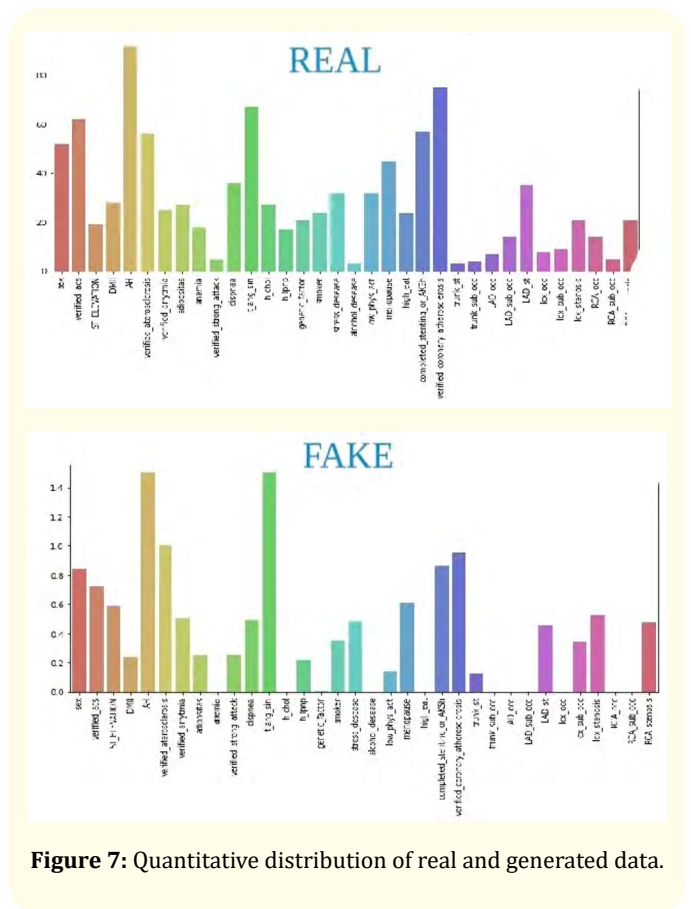


Figure 7: Quantitative distribution of real and generated data.

A heat map was created for comparing basic descriptive statistics (median, mean 25 quantile, 75 quantile, minimum and maximum values). Significant differences were obtained in half of the signs (Figure 8).

A heat map of the difference between the correlation matrices of real and synthetic datasets has been created. The main correlation components are preserved (Figure 9).

Comparing the synthetic data with the real ones, we can conclude that the generated data are close to the real ones. The main basic flow dependencies of features are preserved, however, the generated dataset does not completely copy the dependencies of the real one, so we can conclude that there are new excellent “random” observations.

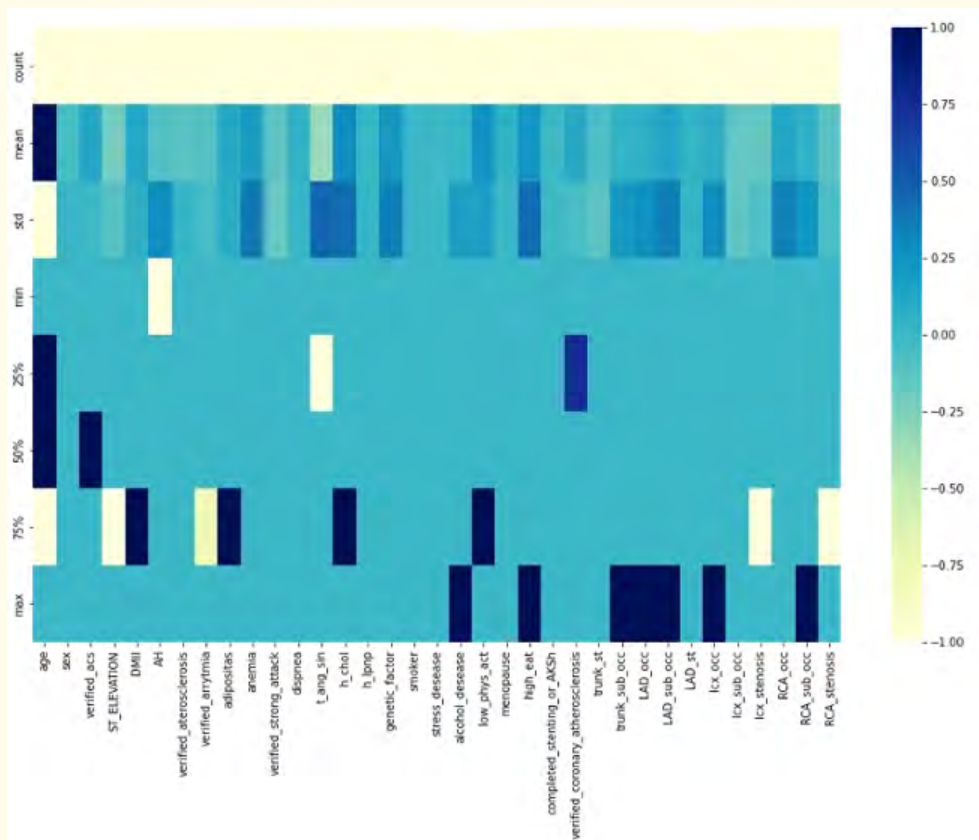


Figure 8: Heat map of the difference in base descriptive statistics.

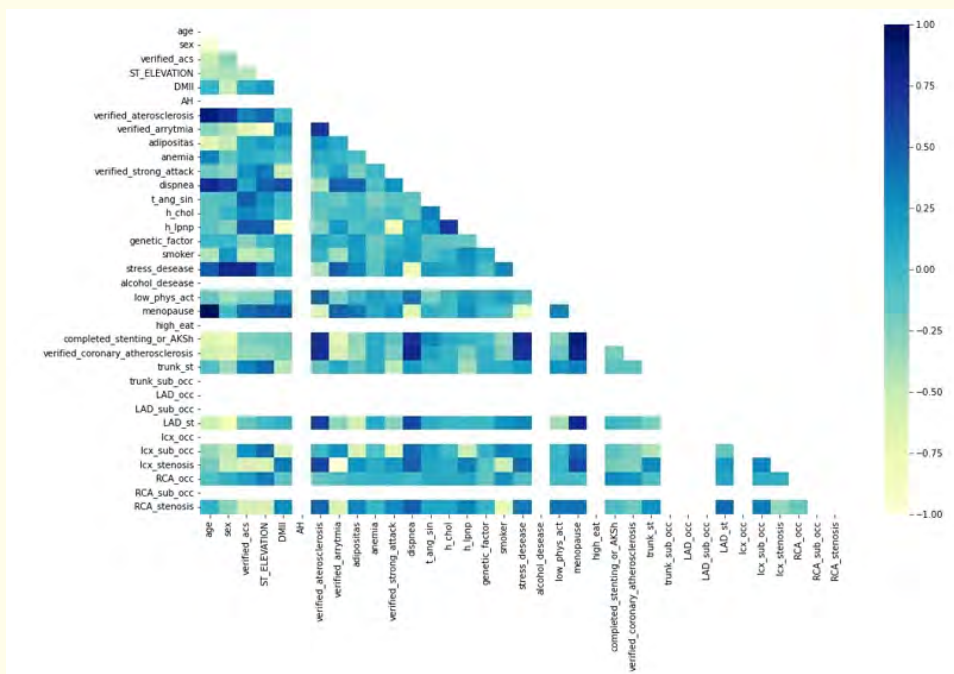


Figure 9: Heat map of the difference in the correlation matrices of the datasets.

## Results and Discussion

### Test sample 1

The prediction of damage to the main coronary arteries and transient myocardial ischemia was carried out.

On a test sample of 20 patients, the result of the neural network was: AUC score 0.74, accuracy (accuracy) reached 80%, “precision” accuracy (precision) — 63%, recall (recall) — 55%, f1 score — 59%.

Average response rates of cardiologists: accuracy 76%, precision 48%, recall 55%, AUC score 0.68, f1 score 49%. The best values among cardiologists were: AUC score 0.72, accuracy 76%, precision 48%, recall 67%, f1 score 56% (Table 1).

	AUC	Accuracy	Precision	Recall	F1 score
Non-invasive predictive AI coronary angiography	74	80	63	55	59
Average answers of specialists	68	76	48	55	49
Best expert answer	72	76	48	67	56

**Table 1:** Prediction of damage to the main coronary arteries and transient myocardial ischemia on test sample 1.

### Test sample 2

The prediction of damage to the main coronary arteries and myocardial ischemia was carried out.

On a test sample of 30 patients, the AUC score was 0.87. Accuracy reached 96%, “precision” accuracy (precision) — 76%, recall (recall) — 71%, f1 score — 74.1% (Table 2).

	AUC	Accuracy	Precision	Recall	F1 score
Non-invasive predictive AI coronary angiography	87	96	76	71	74

**Table 2:** Prediction of damage to the main coronary arteries and transient myocardial ischemia on test sample 2.

The efficiency of detecting myocardial ischemia was compared based on the calculation of the prognosis of the need for coronary artery revascularization by neural network analysis and the results obtained when performing daily ECG monitoring and treadmill test.

Results of the neural network analysis method: accuracy 93%, precision 60%, recall 100%, AUC score 96%, f1 score 75%, daily ECG monitoring: accuracy 87%, precision 33%, recall 33%, AUC score 63%, f1 score 33%, treadmill test: accuracy 70%, precision 12%, recall 33%, AUC score 54%, f1 score 18% (Table 3).

	AUC	Accuracy	Precision	Recall	F1 score
Non-invasive predictive AI coronary angiography	96	93	60	100	75
Daily ECG monitoring	63	87	33	33	33
Treadmill test	54	70	12	33	18

**Table 3:** Detection of myocardial ischemia on test sample 2.

### Test sample 3

The prediction of damage to the main coronary arteries and transient myocardial ischemia for 1,500,000 synthetic observations was carried out.

The AUC score was 0.79. Accuracy reached 88%, “precision” accuracy (precision) — 73%, recall (recall) — 63%, f1 score — 67% (Table 4).



	AUC	Accuracy	Precision	Recall	F1 score
Non-invasive predictive AI coronary angiography	79	88	73	63	67

**Table 4:** Prediction of damage to the main coronary arteries and transient myocardial ischemia on test sample 3.

The created neural network analysis model makes it possible to predict damage to the main coronary arteries with sufficient probability based on structured data and ECG images. The accuracy of detecting transient myocardial ischemia, determined by the method of neural network analysis, obtained in order to predict the need for coronary artery revascularization, is higher than that of classical diagnostic methods, such as daily ECG monitoring and treadmill test. The results obtained allow us to speak about the possible practical application of the neural set analysis method in clinical practice.

The innovative approach lies in the use of neural networks for the diagnosis of coronary artery pathology based on risk factors and ECG images. At the output of the neural network, we get the presence or absence of pathology on each main coronary artery (trunk of the left coronary artery, anterior interventricular artery, circumflex artery, right coronary artery), the likelihood of atherosclerosis, the need to perform invasive coronary angiography with possible revascularization at the moment. The advantage of using our method is simplicity (it requires filling out a questionnaire and uploading an ECG image), speed (calculation time is less than a second), non-invasiveness of the technique while maintaining high accuracy. Our technique can be used remotely and will allow performing non-invasive predictive AI coronary angiography in places where there is no possibility of specialized medical care (removed ECG tape is required). It also does not require extensive computer resources and expensive equipment, which makes it easier for a specialist to make a correct diagnosis. The system helps to identify an acute condition at an early stage by the type of occlusion, subocclusion, significant stenosis of the coronary arteries, which will serve as an early reason to contact a specialist. Immediate results are a huge advantage over other studies that require results to be expected within 24 to 48 hours. The duration of the study takes several minutes, and also does not require the cost of qualified medical labor, which would reduce the workload for doctors. Any hospital employee will be able to ask a few questions to the patient, as well as upload an ECG tape to the system. When introducing the study into

the system of providing medical care to citizens under compulsory medical insurance, we could receive patients with an “increased risk category”. Upon receipt of such a result, it would be possible to register these patients with a cardiologist as a priority and then for the necessary additional studies required for a more accurate diagnosis. Our technique allows us to approach the screening of coronary artery pathology at a new level, the study can be used massively due to the lack of “invasiveness”, the introduction of contrast studies, myocardial overload. Our program will allow you to independently suspect and identify the presence of pathology in a patient. In case of receiving a “positive result”, the patient could immediately make an appointment with a doctor, having an increased risk category. In our work, we tried as much as possible to bring the work of AI closer to the work of a doctor.

Our system processes any ECG images without reference to a specific ECG device. It is enough to take a picture of the ECG and upload it to the system. For example, you can take a picture of an ECG tape on a mobile phone and the system will make a prediction.

An important additional advantage of neural network data analysis is the fact that, when treating chronic coronary syndrome, cardiologists do not have reliable “tools” for an undoubted referral to coronary angiography, and under these conditions, artificial intelligence makes it possible to correctly interpret the data set and direct the doctor to perform interventional technology. It is also worth noting that in senile patients without symptoms, with limited ability to perform stress testing on the fact of chronic coronary syndrome, the deep machine learning technique provides an invaluable prospect for the timely referral of the patient for coronary angiography.

In the future, we plan to conduct studies to assess the accuracy of the neural network model separately for tabular data and ECG images. Thus, it will be possible only on the basis of tabular data or ECG images to obtain a forecast of damage to the main arteries of the heart.

## Conclusion

Neural network analysis of the prepared clinical, laboratory and instrumental data allows you to adjust the network parameters for subsequent prediction of damage to the main coronary arteries. The neural network trained by us predicts damage to the main coronary arteries with a sensitivity of 63%, a specificity of 88%, and AUC of 0.74.

On the test sample, the neural network works more efficiently than the average cardiologists and, what is especially important, allows the doctor to be directed to perform invasive examination methods in cases where there is not enough input data for this decision. One in five experts was able to get close to the accuracy of the trained neural network model. The efficiency of detecting transient myocardial ischemia in a test sample is higher for a trained neural network compared to classical diagnostic methods, such as daily ECG monitoring, treadmill test.

On an extremely large sample of 1500000 observations, a high AUC score was obtained.

## Conflict of Interest

All authors who have taken part in this study declare that they have nothing to disclose.

## Bibliography

1. Kuo FC., *et al.* "The relative utilities of genome-wide, gene panel, and individual gene sequencing in clinical practice". *Blood* 130 (2017): 433-439.
2. Muse ED., *et al.* "Towards a smart medical home". *Lancet* 389 (2017): 358.
3. Steinhubl SR., *et al.* "The emerging field of mobile health". *Science Translational Medicine* 7 (2015): 283rv3.
4. Shameer K., *et al.* "Translational bioinformatics in the era of real-time biomedical, health care and wellness data streams". *Brief Bioinformation* 18 (2017): 105-124.
5. Konstam MA., *et al.* "The academic medical system: reinvention to survive the revolution in health care". *Journal of the American College of Cardiology* 69 (2017): 1305-1312.
6. Steinhubl SR and Topol EJ. "Moving from digitalization to digitization in cardiovascular care: why is it important, and what could it mean for patients and providers?" *Journal of the American College of Cardiology* 66 (2015): 1489-1496.
7. Boeldt DL., *et al.* "How consumers and physicians view new medical technology: comparative survey". *Journal of Medical Internet Research* 17 (2015): e215.
8. Vysotskaya ZhM and Terzov AI. "Mathematical models of non-invasive determination of coronary artery lesions in patients with coronary heart disease". On Sat. New applications of morphometry and mathematical modeling in biomedical research. Kharkov, (1990): 53.
9. Bala YuM., *et al.* "Mathematical approach to automatic diagnosis of ischemic heart disease". On Sat. Computerization in medicine. Voronezh, (1990): 66-70.
10. A Waheed., *et al.* "CovidGAN: Data Augmentation Using Auxiliary Classifier GAN for Improved Covid-19 Detection". in *IEEE Access* 8 (2020): 91916-91923.