

Оригинальная статья

Заполнение пропусков в данных в базе пациентов кардиологического профиля линейными методами при прогнозировании летального исхода

Бибичева Т.С.¹, Сказкина В.В.¹, Огнева М.В.¹, Симонян М.А.², Гриднев В.И.², Караваев А.С.^{1,2,3}¹ФГБОУ ВО Саратовский национальный исследовательский государственный университет имени Н.Г. Чернышевского, Саратов, Россия²ФГБОУ ВО Саратовский государственный медицинский университет им В. И. Разумовского Минздрава России, Саратов, Россия³Саратовский филиал Института радиотехники и электроники им. В.А. Котельникова РАН, Саратов, Россия

Резюме

В работе проведено заполнение пропущенных данных в базе данных федерального регистра острого коронарного синдрома и выполнена оценка вероятности прогнозирования летального исхода. Для заполнения использованы линейные методы замены наиболее вероятным или средним значением. Решение задачи прогнозирования выполнено с использованием метода k-ближайших соседей. Показано, что используемые методы заполнения пропусков, несмотря на свою простоту, повышают вероятность прогнозирования летальных исходов на 6%.

Ключевые слова: регистр, анализ данных, пропущенные данные.

Библиографическая ссылка: Бибичева Т.С., Сказкина В.В., Огнева М.В., Симонян М.А., Гриднев В.И., Караваев А.С. Заполнение пропусков в данных в базе пациентов кардиологического профиля линейными методами при прогнозировании летального исхода. *Кардио-ИТ* 2021; 8(1): e0101.

Поступила в редакцию 22 марта 2021. Принята в печать 2 июня 2021.

© 2021, Бибичева Т.С., Сказкина В.В., Огнева М.В., Симонян М.А., Гриднев В.И., Караваев А.С.

Ответственный автор: Бибичева Татьяна Сергеевна – ФГБОУ ВО СГУ им. Н.Г. Чернышевского, 83, ул. Астраханская, 410012, Саратов, Россия.

Тел.: +7 (937) 974-66-50. E-mail: bibichevatat@gmail.com

Original article

Missing value imputation with linear methods in the database of cardiological patients in prediction of mortality

Tatiana S. Bibicheva¹, Viktoriia V. Skazkina¹, Marina V. Ogneva¹, Margarita A. Simonyan², Vladimir I. Gridnev², Anatoly S. Karavaev^{1,2,3}¹Saratov State University, Saratov, Russia²Saratov State Medical University, Saratov, Russia³Saratov Branch of the Institute of Radio-Engineering and Electronics of Russian Academy of Sciences, Saratov, Russia

Abstract

This study examines missing value imputation in the Russian Acute Coronary Syndrome Registry (RusACSR) and assessment of the probability of predicting mortality. Linear methods with the most probable or average value were used for imputation. The prediction problem was solved using the k-nearest neighbors method. This work reveals that the imputation method, despite their simplicity, increases the probability of prediction of mortality by 6%.

Keywords: registry, data analysis, missing values.

Cite as Bibicheva T.S., Skazkina V.V., Ogneva M.V., Simonyan M.A., Gridnev V.I., Karavaev A.S. Missing value imputation with linear methods in the database of cardiological patients in prediction of mortality. *Cardio-IT* 2021; 8(1): e0101.

Received 22 March 2021. Accepted 2 June 2021.

© 2021, Bibicheva T.S., Skazkina V.V., Ogneva M.V., Simonyan M.A., Gridnev V.I., Karavaev A.S.

Corresponding author: Tatiana S. Bibicheva. Address: Saratov State University, 83, Astrakhanskaya str., 410012, Saratov, Russia. Phone: +79379746650. E-mail: bibichevatat@gmail.com

Введение

Одной из важных областей применения машинного обучения в медицине является кардиология: заболевания сердечно-сосудистой системы являются главной причиной смертности в мире [1]. В силу развития технологий количество медицинских данных активно растет, но такой объем информации становится не под силу обработать одному медицинскому специалисту. Для решения этой проблемы используются методы машинного обучения, которые зарекомендовали себя, как эффективное средство инструментальной поддержки решения врача [2]. Например, в работе [3] авторы рассматривают методы машинного обучения

для раннего прогнозирования остановки сердца. Система раннего предупреждения, основанная на глубоком обучении, показала эффективность в ходе сопоставления с алгоритмом случайного леса и логистической регрессии.

В работе [4] методы машинного обучения использовались для оценки эффективности использования койко-мест в стационарах. В результате алгоритм случайного леса показал лучший результат (80%).

Еще одной проблемой является выявление риска сердечно-сосудистых заболеваний, протекающих бессимптомно. В работе [5] использовались данные более 400 тыс. пациентов для прогнозирования на основе 473

параметров. Модель была построена с использованием алгоритмического инструмента, который автоматически настраивает модель машинного обучения. В результате было показано, что модель повышает точность прогнозирования риска сердечно-сосудистых заболеваний.

Существенной проблемой при обработке больших биомедицинских данных считается наличие пропусков в базах данных. Это может быть связано как с изменениями в протоколах, так и неаккуратным заполнением младшим медперсоналом [6].

Решению проблемы пропусков в данных посвящено значительное количество работ. Например, в статье [7] рассматриваются базы данных отделений интенсивной терапии. Отмечается, что отсутствующие данные могут повлиять на эффективность прогнозирования рисков развития патологии. В статье изучаются две наиболее распространенные стратегии работы с недостающими данными: заполнение и удаление значений. Для решения, какие пропущенные значения следует заполнить, а какие удалить, разработан статистический классификатор на основе нечеткого моделирования. Метод позволил улучшить точность классификации на 10%. Заполнение пропущенных данных изучается в работе [8] по прогнозированию смертности от острого инфаркта миокарда для раннего вмешательства и проведения необходимых процедур. Задачей является заполнение пропусков, а не удаление записей с недостающими значениями. Данные заполняются методами с использованием средних значений, линейной регрессии, а также методами, основанными на моделировании с машинным обучением: метод k-ближайших соседей [9], случайный лес [10] и нейронные сети [11]. Лучшие результаты показали методы, основанные на нейросетевых подходах, и метод случайного леса. Сравнение заполнения лабораторных данных с использованием методов случайного леса, средних значений, ближайшего соседа и линейных методов рассматривается в работе [12]. Метод случайного леса показал наименьшую ошибку, как для непрерывных, так и для категориальных данных, а также наименьшую ошибку прогноза для априорно-известных значений.

Таким образом, методы машинного обучения продемонстрировали свою эффективность при решении задач медицинской диагностики прогнозирования развития патологий, при этом значительной проблемой являются пропуски и неполнота этих данных.

Одной из наиболее известных и крупных российских баз является федеральный регистр острого коронарного синдрома [13]. База содержит более 400 тыс. записей пациентов, при этом каждая запись включает около 200 параметров. Эта база данных позволяет анализировать демографические, клинические, гендерные и другие особенности различных групп пациентов [14]. Однако, несмотря на высокую информативность данной базы, проблемы пропусков в данных и ошибок при их вводе младшим медицинским персоналом остаются существенной проблемой, поэтому целью данной работы является изучение статистики пропусков в этой базе данных и оценка возможности заполнения пропусков линейными методами в ходе прогнозирования исхода лечения при таком заполнении на примере предсказания летальных исходов для пациентов кардиологического профиля.

Описание данных

В работе анализировалась база данных федерального регистра острого коронарного синдрома [13]. Для оценки возможностей метода были выбраны данные за 2014-2018 гг, до начала эпидемии COVID-19 для исключения влияния пандемии на статистику.

Анализируемая выборка включала 163 тыс. записей, каждая запись включает 168 параметров. Параметры делятся на количественные и качественные. К количественным параметрам относятся: возраст, рост, вес, дата выписки-дата поступления, результаты лабораторных анализов уровня глюкозы плазмы крови, гемоглобина и т.п., к качественным: пол, жизненный статус, анамнез, назначенное лечение и пр. Записи анализируемой выборки включали информацию об исходе лечения: пациент жив/мертв при выписке. Из записей были алгоритмически удалены заведомо ошибочные значения. Анализируемая выборка включала записи пациентов, страдающих различными формами ишемической болезни сердца и перенесших инфаркт миокарда.

Средний возраст пациентов составил 65.0 ± 11.8 лет (среднее±стандартное отклонение). Выборка включала 40.1% женщин, средний возраст 70.7 лет, 59.9% мужчин, средний возраст – 61.8 лет. Анализируемые записи имели пропуски в данных.

Методы

На первом этапе проводилась предварительная обработка базы данных с алгоритмической коррекцией заведомо невозможных значений на основе априорной информации. Корректировались параметры, связанные со временем, отрицательные значения считались ошибочными, соответствующие записи были удалены. Таким образом, были обработаны признаки “Дата выписки-Дата поступления”, “Дата возникновения симптомов-Дата вызова скорой помощи” и др. Для параметров, связанных с анализами крови, весом, ростом, была выполнена проверка на выход за априори известные из методических рекомендаций диапазоны значений.

Для заполнения пропусков в данных использовались линейные статистические методы: заполнение средними значениями для количественных параметров и заполнение наиболее часто встречающимся значением для качественных параметров.

Эффективность заполнения пропусков оценивалась в ходе прогнозирования исхода результата лечения пациентов методом k-ближайших соседей. Оценивались вероятности верного предсказания летального исхода и выживания пациентов.

Метод k-ближайших соседей [9] реализуется в многомерном пространстве параметров записей испытуемых. Пусть есть множество объектов X (множество записей анализируемой выборки), множество допустимых ответов Y (в данной задаче параметр Жизненный статус). Существует целевая зависимость y^* : $X \rightarrow Y$, значения которой известны только на объектах обучающей выборки $X^l = \{(x_i, y_i)\}_{i=1}^l$, $y_i = y^*(x_i)$, где l – объем обучающей выборки, (x_i, y_i) – пары «объект-ответ», y^* – целевая функция. Необходимо построить алгоритм a для классификации: $a: X \rightarrow Y$, аппроксимирующий целевую зависимость $y^*(x)$ на всем множестве X .

Алгоритм k-ближайших соседей относит классифицируемый объект $x \in X^l$ к тому классу, элементов

которого окажется больше среди k -ближайших соседей $x_{u^k(i)}, i=1, \dots, k$.

Результаты

Первым этапом работы с данными являлась коррекция заведомо невозможных значений. Было выявлено 18143 таких значений в 14871 (8.3%) записях. Наиболее часто корректируемым параметром оказался параметр “Дата проведения тромболитического-Дата госпитализации”, таких записей выявлено 4.9%.

Следующим шагом был анализ пропусков. Всего в экспериментальной выборке имеется 8103923 (29.6%) пропущенных значений параметров, в среднем 49 пропусков для каждой записи. Наиболее вероятны пропуски в параметрах: “Дата возникновения симптомов - дата вызова скорой помощи” – 54.4%, “Дата прибытия скорой помощи - Дата госпитализации (минуты)” – 58%, “Семейная история раннего развития ИБС” 44.3%, результатах лабораторных исследований, например, “Диагностически значимое повышение тропонина I”, 56.2%, назначенном лечении “Антагонисты кальция недигидропиридиновые” 42.7%.

При заполнении пропущенных значений некоторые параметры были определены заранее, например, параметр “Число установленных стентов” заполнялся значением 0 при отсутствии данных.

Далее для восстановления использовались метод заполнения средними значениями для количественных параметров и заполнение наиболее часто встречающимся значением для качественных параметров.

Чтобы оценить результаты заполнения, была проведена оценка вероятности прогнозирования исхода лечения на примере предсказания летальных исходов для базы данных, заполненной нулями, и базы, заполненной описанным способом. Для решения задачи прогнозирования использовался метод k -ближайших соседей.

При использовании методов машинного обучения необходима настройка гиперпараметров методов, для метода k -ближайших соседей основной настраиваемый гиперпараметр k – количество соседей, используемое при обучении. При работе был выполнен подбор количества из списка [3, 5, 7, 9]. Для задачи с двумя классами количество соседей k выбирается нечетным, т.к. максимум может достигаться сразу на обоих классах. Для оценки подбора оценивается точность прогнозирования, метод с гиперпараметром $k=5$ показал лучший результат для выборки данных.

Для обучения метода выборка делилась на обучающую и тестовую. Для обучающей выборки случайным образом выбраны 60% записей, для тестовой выборки – 40% записей.

Всего в выборке было 157650 человек с летальным исходом и 5405 выживших. Классы являются несбалансированными, результаты представлены отдельно для каждого класса.

При заполнении нулями точность прогнозирования летального исхода составила 0.74, выживания – 0.94, при заполнении средними значениями и наиболее часто встречающимися – точность прогнозирования летального исхода равна 0.80, для выживания – 0.96.

Заключение

Наличие пропусков в больших данных является проблемой [15,16]. Анализ базы данных федерального регистра острого коронарного синдрома показал, что проблема ошибочно введенных данных и пропусков является существенной и в анализируемой выборке эти проблемы встречаются во всех записях. Причинами этого могут являться технические сбои, недостатки правового регулирования, которое не регламентирует требования по заполнению баз данных медицинским персоналом.

Административно-организационные вопросы выходят за рамки этой работы, но технические приемы позволяют снизить остроту проблемы.

Была проведена оценка вероятности наличия пропусков в базе данных федерального регистра острого коронарного синдрома, пропуски в данных были заполнены с помощью метода замены часто встречающимся или средним значением, затем проведена оценка вероятности прогнозирования исхода лечения на примере предсказания летальных исходов для пациентов кардиологического профиля. Вероятность верного предсказания летального исхода после заполнения пропусков возросла на 6%, вероятность предсказания выживаемости возросла на 2%.

В данной работе были использованы простые линейные методы заполнения пропусков. Результаты их применения позволяют предположить, что применение более сложных нелинейных методов машинного обучения может улучшить результаты прогнозирования.

Литература

1. World Health Organization. Cardiovascular diseases (CVDs). <https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-cvds>
2. Yoruk U., Hargreaves B.A., Vasanawala S.S. Automatic renal segmentation for MR urography using 3D-GrabCut and random forests. *Magn Reson Med*. 2018; 79(3): 1696-1707. <http://dx.doi.org/10.1002/mrm.26806>.
3. Kwon J.M., Lee Y., Lee S., et al. An Algorithm Based on Deep Learning for Predicting In-Hospital Cardiac Arrest. *J Am Heart Assoc*. 2018; 7(13): e008678. <http://dx.doi.org/10.1161/JAHA.118.008678>.
4. Daghastani T.A., Elshawi R., Sakr S., et al. Predictors of in-hospital length of stay among cardiac patients: A machine learning approach. *Int J Cardiol*. 2019; 288: 140-147. <http://dx.doi.org/10.1016/j.ijcard.2019.01.046>.
5. Alaa A.M., Bolton T., Di Angelantonio E., et al. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. *PLoS One*. 2019; 14(5): e0213653. <http://dx.doi.org/10.1371/journal.pone.0213653>.
6. Тихова Г.А. Пропуск данных в выборке: как решать проблему и как ее избежать. *Региональная анестезия и лечение острой боли* 2016; 10 (3): 205-209.
7. Cismondì F., Fialho A.S., Vieira S.M., et al. Missing data in medical databases: impute, delete or classify? *Artif Intell Med* 2013; 58(1): 63-72. <http://dx.doi.org/10.1016/j.artmed.2013.01.003>.
8. Payrovnaziri S.N., Xing A., Salman S., et al. The Impact of Missing Value Imputation on the Interpretations of Predictive Models: A Case Study on One-year Mortality Prediction in ICU Patients with Acute Myocardial Infarction. *MedRxiv* 2020. <http://dx.doi.org/10.1101/2020.06.06.20124347>.
9. Воронцов К.В. Математические методы обучения по прецедентам. 2007; 141 с.

10. Müller A., Guido Sarah. Introduction to Machine Learning with Python: A Guide for Data Scientists. *O'Reilly Media* 2017; 400 p.
11. Хайкин С. Нейронные сети: полный курс. М: «Вильямс», 2006; 1104 с.
12. Waljee A.K., Mukherjee A., Singal A.G., et al. Comparison of imputation methods for missing laboratory data in medicine. *BMJ Open* 2013; 3: e002847. <http://dx.doi.org/10.1136/bmjopen-2013-002847>
13. Ощепкова Е.В., Довгалецкий П.Я., Гриднев В.И., и др. Руководство пользователя информационно-аналитической системы «Федеральный регистр больных с острым коронарным синдромом». *Кардио-ИТ* 2014; 1: 0203. <https://doi.org/10.15275/cardioit.2014.0203>.
14. Gridnev, V.I., Kiselev, A.R., Posnenkova, et al. Objectives and Design of the Russian Acute Coronary Syndrome Registry (RusACSR). *Clin Cardiol* 2016; 39: 1-8. <https://doi.org/10.1002/clc.22495>
15. Hu Z., Melton G.B., Arsoniadis E.G., et al. Strategies for handling missing clinical data for automated surgical site infection detection from the electronic health record. *J Biomed Inform* 2017; 68: 112-120. <http://dx.doi.org/10.1016/j.jbi.2017.03.009>.
16. Phung S., Kumar A., Kim J. A deep learning technique for imputing missing healthcare data. *Annu Int Conf IEEE Eng Med Biol Soc.* 2019: 6513-6516. <http://dx.doi.org/10.1109/EMBC.2019.8856760>.
13. Oschepkova EV, Dovgalevsky PY, Gridnev VI, et al. User manual of the information and analytical system Russian Acute Coronary Syndrome Registry. *Cardio-IT* 2014; 1: 0203 [In Russ]. <https://doi.org/10.15275/cardioit.2014.0203>
14. Gridnev VI, Kiselev AR, Posnenkova OM, et al. Objectives and Design of the Russian Acute Coronary Syndrome Registry (RusACSR). *Clin Cardiol* 2016; 39: 1-8. <https://doi.org/10.1002/clc.22495>.
15. Hu Z, Melton GB, Arsoniadis EG, et al. Strategies for handling missing clinical data for automated surgical site infection detection from the electronic health record. *J Biomed Inform* 2017; 68: 112-120. <http://dx.doi.org/10.1016/j.jbi.2017.03.009>.
16. Phung S, Kumar A, Kim J. A deep learning technique for imputing missing healthcare data. *Annu Int Conf IEEE Eng Med Biol Soc.* 2019: 6513-6516. <http://dx.doi.org/10.1109/EMBC.2019.8856760>.

References

1. World Health Organization. Cardiovascular diseases (CVDs). <https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-cvds>.
2. Yoruk U, Hargreaves BA, Vasanawala SS. Automatic renal segmentation for MR urography using 3D-GrabCut and random forests. *Magn Reson Med.* 2018; 79(3): 1696-1707. <http://dx.doi.org/10.1002/mrm.26806>.
3. Kwon JM, Lee Y, Lee S, et al. An Algorithm Based on Deep Learning for Predicting In-Hospital Cardiac Arrest. *J Am Heart Assoc.* 2018; 7(13): e008678. <http://dx.doi.org/10.1161/JAHA.118.008678>.
4. Daghistani TA, Elshawi R, Sakr S, et al. Predictors of in-hospital length of stay among cardiac patients: A machine learning approach. *Int J Cardiol.* 2019; 288: 140-147. <http://dx.doi.org/10.1016/j.ijcard.2019.01.046>.
5. Alaa AM, Bolton T, Di Angelantonio E, et al. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. *PLoS One.* 2019; 14 (5): e0213653. <http://dx.doi.org/10.1371/journal.pone.0213653>.
6. Tikhova GA. Missing data in the set: how to solve the problem and how to avoid it. *Regional anesthesia and therapy of acute pain* 2016; 10 (3): 205-209 [In Russ].
7. Cisondi F, Fialho AS, Vieira SM, et al. Missing data in medical databases: impute, delete or classify? *Artif Intell Med* 2013; 58(1): 63-72. <http://dx.doi.org/10.1016/j.artmed.2013.01.003>.
8. Payrovnaziri SN, Xing A, Salman S, et al. The Impact of Missing Value Imputation on the Interpretations of Predictive Models: A Case Study on One-year Mortality Prediction in ICU Patients with Acute Myocardial Infarction. *MedRxiv* 2020. <http://dx.doi.org/10.1101/2020.06.06.20124347>.
9. Vorontsov KV. Mathematical teaching methods by precedents. 2007; 141 p. [In Russ].
10. Müller A, Guido Sarah. Introduction to Machine Learning with Python: A Guide for Data Scientists. *O'Reilly Media* 2017; 400 p.
11. Haykin S. Neural Networks: A Comprehensive Foundation. Moscow: Williams 2006; 1104 p. [In Russ].
12. Waljee AK, Mukherjee A, Singal AG, et al. Comparison of imputation methods for missing laboratory data in medicine. *BMJ Open* 2013; 3: e002847. <http://dx.doi.org/10.1136/bmjopen-2013-002847>.
13. Сказкина Виктория Викторовна – аспирант, кафедра динамического моделирования и биомедицинской инженерии, ФГБОУ ВО СГУ им. Н.Г. Чернышевского, Саратов, Россия. <https://orcid.org/0000-0001-7860-2616>.
14. Огнева Марина Валентиновна – канд. физ.-мат. наук, доцент кафедры информатики и программирования, ФГБОУ ВО СГУ им. Н.Г. Чернышевского, Саратов, Россия. <http://orcid.org/0000-0002-9828-7681>
15. Симонян Маргарита Андреевна – врач-кардиолог, младший научный сотрудник отдела атеросклероза и хронической ишемической болезни сердца, НИИ кардиологии, ФГБОУ ВО Саратовский ГМУ им. В.И. Разумовского Минздрава России, Саратов, Россия. <http://orcid.org/0000-0002-9866-3069>.
16. Гриднев Владимир Иванович – д-р мед. наук, доцент, директор НИИ кардиологии, ФГБОУ ВО Саратовский ГМУ им. В.И. Разумовского Минздрава России, Саратов, Россия. <http://orcid.org/0000-0001-6807-7934>.
17. Караваев Анатолий Сергеевич – д.ф.-м.н., профессор кафедры динамического моделирования и биомедицинской инженерии, ФГБОУ ВО СГУ им. Н.Г. Чернышевского, Саратов, Россия; ведущий научный сотрудник ФГБОУ ВО Саратовский ГМУ им. В.И. Разумовского Минздрава России, Саратов, Россия; старший научный сотрудник лаборатории моделирования в нелинейной динамике Саратовского филиала Института радиотехники и электроники им. В.А. Котельникова РАН, Саратов, Россия. <http://orcid.org/0000-0003-4678-3648>.

Информация об авторах:

Бибичева Татьяна Сергеевна – аспирант, кафедра динамического моделирования и биомедицинской инженерии, ФГБОУ ВО СГУ им. Н.Г. Чернышевского, Саратов, Россия. <https://orcid.org/0000-0001-7860-2616>.

Сказкина Виктория Викторовна – аспирант, кафедра динамического моделирования и биомедицинской инженерии, ФГБОУ ВО СГУ им. Н.Г. Чернышевского, Саратов, Россия. <http://orcid.org/0000-0001-9380-8292>.

Огнева Марина Валентиновна – канд. физ.-мат. наук, доцент кафедры информатики и программирования, ФГБОУ ВО СГУ им. Н.Г. Чернышевского, Саратов, Россия. <http://orcid.org/0000-0002-9828-7681>

Симонян Маргарита Андреевна – врач-кардиолог, младший научный сотрудник отдела атеросклероза и хронической ишемической болезни сердца, НИИ кардиологии, ФГБОУ ВО Саратовский ГМУ им. В.И. Разумовского Минздрава России, Саратов, Россия. <http://orcid.org/0000-0002-9866-3069>.

Гриднев Владимир Иванович – д-р мед. наук, доцент, директор НИИ кардиологии, ФГБОУ ВО Саратовский ГМУ им. В.И. Разумовского Минздрава России, Саратов, Россия. <http://orcid.org/0000-0001-6807-7934>.

Караваев Анатолий Сергеевич – д.ф.-м.н., профессор кафедры динамического моделирования и биомедицинской инженерии, ФГБОУ ВО СГУ им. Н.Г. Чернышевского, Саратов, Россия; ведущий научный сотрудник ФГБОУ ВО Саратовский ГМУ им. В.И. Разумовского Минздрава России, Саратов, Россия; старший научный сотрудник лаборатории моделирования в нелинейной динамике Саратовского филиала Института радиотехники и электроники им. В.А. Котельникова РАН, Саратов, Россия. <http://orcid.org/0000-0003-4678-3648>.

Authors:

Tatiana S. Bibicheva – PhD student, MSc, Department of Dynamic Modeling and Biomedical Engineering, Saratov State University, Saratov, Russia. <https://orcid.org/0000-0001-7860-2616>.

Viktoriia V. Skazkina – PhD student, MSc, Engineer, Department of Dynamic Modeling and Biomedical Engineering, Saratov State University, Saratov, Russia. <http://orcid.org/0000-0001-9380-8292>.

Marina V. Ogneva – PhD, Assistant Professor, Department Chair, Department of Informatics and Programming, Saratov State University, Saratov, Russia. <http://orcid.org/0000-0002-9828-7681>.

Margarita A. Simonyan – MD, Junior Researcher of Department of Atherosclerosis and Chronic Ischemic Heart Disease of Research Institute of Cardiology, Saratov State Medical University, Saratov, Russia. <http://orcid.org/0000-0002-9866-3069>.

Vladimir I. Gridnev – MD, DSc, Director of Research Institute of Cardiology, Saratov State Medical University, Saratov, Russia. <http://orcid.org/0000-0001-6807-7934>.

Anatoly S. Karavaev – DSc, Professor, Department of Dynamic Modeling and Biomedical Engineering, Saratov State University, Saratov, Russia; Leading Researcher, Department of Innovative Cardiological Information Technology, Institute of Cardiological Research, Saratov State Medical University, Saratov, Russia; Senior Researcher, Laboratory of Nonlinear Dynamics Modelling, Saratov Branch of the Institute of RadioEngineering and Electronics of Russian Academy of Sciences, Saratov, Russia. <http://orcid.org/0000-0003-4678-3648>.