

**МАТЕМАТИЧЕСКИЕ МОДЕЛИ ПРОГНОЗИРОВАНИЯ
ТЕЧЕНИЯ ФИБРИЛЛЯЦИИ ПРЕДСЕРДИЙ
ПРИ БЕРЕМЕННОСТИ¹**

Н. Е. Дятлов, Н. Ю. Митрохина, Ф. К. Рахматуллов, М. А. Митрохин

**MATHEMATICAL MODELS FORECASTING
OF ATRIAL FIBRILLATION DURING PREGNANCY**

N. E. Dyatlov, N. Yu. Mitrokhina, F. K. Rakhmatullov, M. A. Mitrokhin

Аннотация. *Предмет и цель работы.* Приоритетной задачей современной кардиологии стоит снижение бремени фибрилляции предсердий (ФП) в популяции. Существующие стратификационные шкалы ФП преследуют цель оценить частоту развития инсульта и кровотечений, в то время как шкала для оценки прогрессирования самой ФП до сих пор не разработана. Цель исследования – разработать математические модели прогнозирования количества и длительности пароксизмов ФП путем отбора целевых предикторов развития аритмии на разных сроках беременности. *Материал и методы.* Для построения математических моделей предикции ФП была использована выборка из записей суточного мониторирования ЭКГ (СМ) 90 беременных женщин с пароксизмальной изолированной ФП на каждом сроке беременности. Прогнозирование выполнялось с использованием корреляционного и регрессионного анализа с распределенным лагом. *Результаты.* Корреляционный анализ Спирмена и анализ окрестности соседних компонент (NCA) позволили отобрать 4 специфических инструментальных предиктора развития ФП при беременности из генеральной совокупности в 30 признаков: средняя ночная частота сердечных сокращений, суточное число наджелудочковых и желудочковых экстрасистол, рефрактерный период атриовентрикулярного узла и предсердий. Методом регрессионного анализа построены четыре предикционные модели течения ФП при беременности с использованием указанных предикторов. Проверка на ошибки прогнозирования выявила высокую точность выбранных моделей. *Заключение.* Были разработаны математические модели прогнозирования течения ФП у беременных женщин, которые позволяют врачам-клиницистам и научным работникам прогнозировать динамику заболевания в разные сроки гестации.

Ключевые слова: фибрилляция предсердий, беременность, регрессионный анализ, математическое моделирование.

Abstract. *Subject and goals.* The priority area of modern cardiology is to reduce the burden of atrial fibrillation (AF) in population. The existing AF stratification scales are aimed at assessing the incidence of stroke and bleeding, while the scale for assessing the progression of AF has not yet been developed. The purpose of the study was to identify predictors of AF development in pregnancy and to develop a method for predicting the number and duration of AF paroxysms during different gestational terms. *Materials and methods.* To build mathematical models of the prediction of AF, a sample of daily Holter ECG monitoring (HM) records of 90 pregnant women with paroxysmal isolated AF recorded at each gestational age was used. Prediction was performed using correlation and regres-

¹ Статья подготовлена при поддержке конкурса «Ректорские гранты» в рамках договора №17/19-НИР от 09.04.2019.

sion analysis with distributed lag. *Results.* Spearman's correlation analysis (SCA) and the neighboring components analysis (NCA) allowed us to select 4 specific intrinsic predictors of the development of AF in pregnancy from the general population in 30 signs: mean night-time heart rate (HRMN), total daily number of supraventricular and ventricular extrasystoles (TE), effective refractory period of the atrioventricular node (AVNRP) and atria (ATRP). Using the method of regression analysis, four predictive models of the course of AF during pregnancy were constructed using the specified predictors. *Conclusion.* We have developed mathematical models for predicting the course of AF in pregnant women, which allow clinicians and scientists to guess the dynamics of the disease at different periods of gestation.

Keywords: atrial fibrillation, pregnancy, regression analysis, mathematical modeling.

Введение

Вопросы ведения пациентов с ФП подробно освещены в зарубежных и отечественных клинических рекомендациях [1, 2]. В рекомендациях ВНОА, РКО, ОШК, EHRA, ESC, ACC по диагностике и лечению ФП описаны этиология, патогенез, клиника, диагностика и лечение этого заболевания. Большое внимание уделяется вопросам профилактики развития осложнений ФП – шкала прогнозирования риска инсульта CHA2DS2VASc и шкала риска кровотечения при приеме антикоагулянтов HAS-BLED широко используются в медицинской практике. В то же время методы предупреждения и прогнозирования самой ФП остаются нераскрытыми. Существующий пробел в стратификации динамики развития заболевания в общей популяции открывает широкое поле для фундаментальных научных исследований. Изучение особенностей течения пароксизмов ФП у беременных женщин и триггерных факторов возникновения аритмии в этот период позволяет составить предиктивную модель развития аритмии на любом сроке гестации по исходным показателям [3]. Трудность разработки модели заключается в нескольких моментах: работа с биологическими данными, не поддающимися строгой классификации; необходимость привлечения врача-специалиста; трудность этического взаимодействия с беременной пациенткой. Необходимость оценки и сопоставления данных во времени, т.е. – на разных сроках беременности, ограничивает объем возможно применяемых методик анализа данных. Полученная же модель позволит врачу прогнозировать поведение аритмии и выявлять критические этапы учащения количества и длительности пароксизмов ФП в период гестации, которые могут быть ассоциированы с повышенным риском мозгового инсульта и системных эмболий [4].

Цель настоящего исследования – разработать метод прогнозирования количества и длительности пароксизмов ФП на основании корреляционного и регрессионного анализа путем отбора целевых предикторов развития аритмии на разных сроках беременности.

Материал и методы

Проведенное исследование является проспективным, диагностическим. Оно выполнено в Пензенском государственном университете в рамках междисциплинарной работы врачей и специалистов в области системного анализа, управления и обработки информации. В период с 2014 по 2018 г. обследо-

вано 90 беременных женщин с пароксизмальной изолированной ФП. Медианный возраст обследуемых составил 28 (25; 31) лет. Всем обследуемым выполнено суточное мониторирование ЭКГ (СМ) на малом сроке беременности (до 8 недель), а затем в каждом триместре и спустя 6 месяцев после родов (всего 5 раз), по результатам которого была составлена база данных электрофизиологических маркеров развития ФП на каждом сроке.

Статистические методы включали в себя описательную статистику, углубленный корреляционный и регрессионный анализ. Корреляционный анализ проводили методами Спирмена и анализа окрестности соседних компонент. Для построения математических предиктивных моделей была использована регрессия с распределенным лагом. Валидизация моделей выполнена путем визуального анализа распределения остатков регрессии, расчетом критерия хи-квадрат, корреляционным анализом остатков регрессии, дисперсионным анализом Фишера. Работа выполнена с использованием ПО Microsoft Excel, Microsoft Word, StatSoft Statistica, MathWorks MatLab.

Результаты исследования и их обсуждение

Выделение ключевых триггеров изолированной ФП проводилось в несколько этапов. Первый этап – формирование выборки пациентов и ее первичный анализ для установления общих закономерностей течения аритмии на разных сроках беременности. На втором этапе осуществлялся детальный отбор наиболее сильных коррелятов ФП на каждом сроке беременности и сопоставление их между собой для выделения ключевых триггеров аритмии. На третьем этапе выполнялись построение и идентификация параметров математических моделей прогрессирования ФП с использованием ключевых коррелятов аритмии. Заключительный этап включал проверку качества разработанных моделей.

На первом этапе с помощью анализа СМ были получены следующие электрофизиологические показатели проводящей системы беременных с ФП:

- частотные характеристики синусового ритма: частота сердечных сокращений (ЧСС) – среднесуточная, среднедневная, средненочная, минимальная, максимальная;

- параметры электрофизиологии сердца – время восстановления функции синусового узла и его скорректированный к ЧСС аналог (КВВФСУ); рефрактерный период атриовентрикулярного узла (РПАВ) и предсердий (РПП);

- аритмические характеристики – частота встречаемости наджелудочковых (НЭ) и желудочковых (ЖЭ) экстрасистол за сутки, за день и за ночь.

Анализ жалоб и дневников больных позволил собрать данные по количеству эпизодов ФП за прошедшие три месяца и средней продолжительности одного эпизода ФП, которые были выбраны в качестве искомых величин.

Всего по данным ХМЭКГ было получено свыше 30 инструментальных предикторов возникновения ФП. Работа с моделью прогнозирования, построенной на основе такого количества независимых переменных, для практикующего врача может быть довольно затруднительной. Основным недостатком такой модели является сложность интерпретации полученных с ее помощью результатов.

Повысить интерпретируемость модели возможно путем сокращения входящих в нее параметров. При этом возникает вторая проблема – при со-

кращении параметров модели уменьшается точность получаемых результатов. Таким образом, необходимо было разработать методику отбора предикторов, которая позволила бы отбирать наиболее релевантные решаемой задаче признаки для повышения интерпретируемости модели при незначительном снижении ее точности.

Такая методика может включать два последовательных шага. Первый – предварительный отсев признаков, которые попарно имеют сильную линейную зависимость. Из регрессионного анализа [5] известно, что такие признаки приводят к неустойчивости параметров модели прогнозирования. Поэтому на этом шаге с помощью метода Спирмена (Spearman's correlation analysis – SCA) отбираются из общего массива признаков наиболее зависимые предикторы с коэффициентом корреляции $r > 0,8$. Из каждой пары таких зависимых признаков оставляются по одному.

Второй шаг методики отбора заключается в применении к оставшемуся набору итерационной процедуры анализа «соседних» компонент (neighborhood component analysis – NCA) [6]. Эта процедура реализует непараметрический подбор весов признаков в процессе минимизации функции потерь полного скользящего контроля по отдельным объектам (leave one out – LOO):

$$L(w) = \frac{1}{N} \sum_{i=1}^N l_i + \lambda \sum_{m=1}^M w_m^2,$$

где N – количество априорных примеров в обучающем наборе $S = \{(\mathbf{x}_i, y_i), i = 1, 2, \dots, N\}$; $\mathbf{x}_i \in \mathcal{X}^M$ – вектор признаков; y_i – целевая переменная; l_i – функция потерь регрессии k -ближайших соседей; λ – коэффициент регуляризации.

Подобранные веса отражают значимость признаков при построении модели зависимости целевой переменной [7] и позволяют их ранжировать.

Таким образом, в результате применения описанной методики на втором этапе анализа сначала количество признаков было уменьшено с 30 до 15.

Впоследствии к набору оставшихся 15 предикторов был применен метод NCA. Итогом работы итерационной процедуры NCA стал отбор четырех наиболее значимых признаков для прогноза ФП: среднечасная ЧСС, суточное число НЭ и ЖЭ, РПАВ и РПП. Ввиду сложности расчета последних двух признаков в клинике они были выделены в отдельную группу электрофизиологических предикторов ФП, в то время как два первых признака использовались для построения модели на основе частотных и аритмических характеристик пациента при автоматическом анализе записи СМ.

На третьем этапе проводилась разработка математических моделей прогнозирования ФП. Для этого использовался регрессионный анализ в виде модели с распределенным лагом, при которой в модель включаются значения признаков как в текущий временной период, так и в предшествующий отрезок времени (в нашем случае – в предшествующий триместр беременности).

Полученные нами модели представлены ниже:

$$AFAM = 3,3379 + 0,0256 \times HRMN + 0,0066 \times TE;$$

$$AFAM = 26,6427 + 0,022 \times AVNRP + 0,0213 \times ATRP;$$

$$AFDUR = 4,0657 + 0,0809 \times HRMN + 0,005 \times TE;$$

$$AFDUR = 21,9269 + 0,0196 \times AVNRP + 0,027 \times ATRP ,$$

где $AFAM$ – количество эпизодов ФП за три месяца (первый прогнозируемый признак), $AFDUR$ – средняя продолжительность одного эпизода ФП (второй прогнозируемый признак), $HRMN$ – средненочная ЧСС, TE – суточное число НЭ и ЖЭ, $AVNRP$ – РПАВ, $ATRP$ – РПП.

Четвертый этап заключался в проверке качества разработанных математических моделей. Сначала остатки регрессии $e = y - \hat{y}$ (где e – регрессионный остаток; y – исходная величина; $\hat{y}$ – прогнозируемое значение y) исследовались на нормальность и отсутствие корреляционных зависимостей.

В качестве примера показана гистограмма распределения остатков регрессии при прогнозировании количества пароксизмов ФП на тестовой выборке (рис. 1). Для сравнения, на этом же графике приведена теоретическая плотность распределения нормального закона с параметрами $\mu = 1,0237 \cdot 10^{-15}$ и $\sigma = 0,5173$.

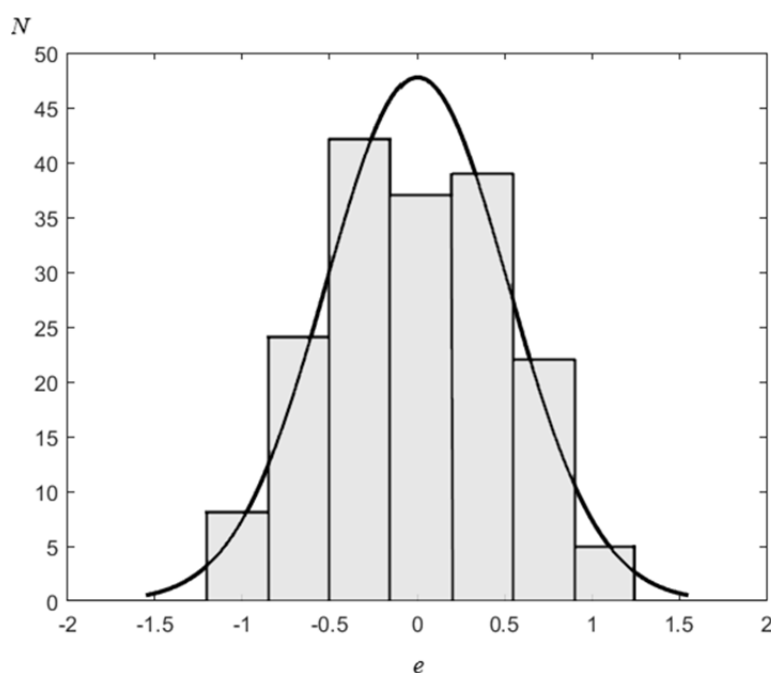


Рис. 1. Гауссовское распределение остатков регрессии

Из рисунка видно, что полученная гистограмма достаточно хорошо аппроксимируется нормальным распределением.

Проверка на нормальность распределения осуществлялась с использованием критерия хи-квадрат: $\chi^2 = 6,3$ ($P < 0,01$).

Отсутствие корреляции между остатками и прогнозными значениями проверялось путем сравнения интересующих коэффициентов корреляции с критическими значениями коэффициентов корреляции, вычисленных через квантили распределения Стьюдента [8]. Факт отсутствия корреляций подтвердился.

Дисперсионный анализ Фишера показал высокую надежность прогностических моделей – критерии рассеяния F для всех моделей оказались статистически выше табличных значений ($P < 0,001$), что говорит о целесообразности включения представленных корреляторов в регрессионные модели.

Заключение

Многофакторный корреляционный и регрессионный анализ данных с учетом их привязки к конкретному временному отрезку позволяет разработать прогнозные модели поведения аритмий в будущем по априорным данным текущего и прошлого моментов времени. Обнаруженные нами в ходе данного исследования инструментальные предикторы возникновения и учащения ФП при беременности легли в основу разработки математических моделей предсказания количества и длительности пароксизмов аритмии в любой период гестации. Полученные модели характеризуются высокой точностью и простотой применения ввиду сравнительно небольшого количества используемых переменных и могут быть рекомендованы к клиническому применению. Предлагаемый научный метод отбора и классификации признаков требует от медицинского работника углубленных знаний в области статистики и системного анализа данных, но может существенно расширить возможности медицинской профилактики и повысить качество оказываемой помощи.

Библиографический список

1. Ревитшвили, А. Ш. Диагностика и лечение фибрилляции предсердий. Рекомендации РКО, ВНОА и АССХ / А. Ш. Ревитшвили. – Москва, 2017. – URL: <http://webmed.irkutsk.ru/doc/pdf/af.pdf> (дата обращения: 14.12.2018).
2. Integrating new approaches to atrial fibrillation management: the 6th AFNET/EHRA Consensus Conference / D. Kotecha, G. Breithardt, A. J. Camm [et al.] // *Europace*. – 2018. – Vol. 1; 20 (3). – P. 395–407.
3. Investigation of models for prognosis of critical values of non-invasive electrophysiological parameters of pregnant women with abnormalities of heart rate / M. Mitrokhin, A. Kuzmin, N. Dyatlov [et al.] // *Conference of Open Innovation Association, FRUCT*. – 2018. – Vol. 21. – P. 238–243.
4. Sauv , N. Atrial Fibrillation in a Structurally Normal Heart during Pregnancy: A Review of Cases From a Registry and From the Literature / N. Sauv ,  . Rey, A. Cumyn // *J Obstet Gynaecol Can.* – 2017. – Vol. 39 (1). – P. 18–24.
5. Айвазян, С. А. Прикладная статистика: Исследование зависимостей : справ. / С. А. Айвазян, И. С. Енюков, Л. Д. Мешалкин ; под ред. С. А. Айвазяна. – Москва : Финансы и статистика, 1985. – 487 с.
6. Neighbourhood components analysis / J. Goldberger, S. Roweis, G. Hinton [et al.] // *Proceedings of the 17th International Conference on Neural Information Processing Systems*. – Canada, 2004. – P. 513–520.
7. Performance comparison of machine learning techniques in sleep scoring based on wavelet features and neighboring component analysis / B. Alizadeh Savareh, A. Bashiri, A. Behmanesh [et al.] // *PeerJ*. – 2018. – e5247.
8. Ферстер, Э. Методы корреляционного и регрессионного анализа. Руководство для экономистов / Э. Ферстер, Б. Ренц ; пер. с нем. и предисловие В. М. Ивановой. – Москва : Финансы и статистика, 1983. – 304 с.

References

1. Revishvili A. Sh. *Diagnostika i lechenie fibrillyatsii predserdiy. Rekomendatsii RKO, VNOA i ASSKh* [Diagnosis and treatment of atrial fibrillation. Recommendations RKO, VNOA and ASH]. Moscow, 2017. Available at: <http://webmed.irkutsk.ru/doc/pdf/af.pdf> (accessed Dec. 14, 2018). [In Russian]
2. Kotecha D., Breithardt G., Camm A. J. et al. *Europace*. 2018, vol. 1; 20 (3), pp. 395–407.
3. Mitrokhin M., Kuzmin A., Dyatlov N. et al. *Conference of Open Innovation Association, FRUCT*. 2018, vol. 21, pp. 238–243.
4. Sauvé N., Rey É., Cumyn A. *J Obstet Gynaecol Can.* 2017, vol. 39 (1), pp. 18–24.
5. Ayvazyan S. A., Enyukov I. S., Meshalkin L. D. *Prikladnaya statistika: Issledovanie zavisimostey: sprav.* [Applied statistics: Research of dependences : handbook]. Moscow: Finansy i statistika, 1985, 487 p. [In Russian]
6. Goldberger J., Roweis S., Hinton G. et al. *Proceedings of the 17th International Conference on Neural Information Processing Systems*. Canada, 2004, pp. 513–520.
7. Alizadeh Savareh B., Bashiri A., Behmanesh A. et al. *PeerJ*. 2018, e5247.
8. Ferster E., Rents B. *Metody korrelyatsionnogo i regressionnogo analiza. Rukovodstvo dlya ekonomistov* [Methods of correlation and regression analysis. A guide for economists]; transl. from Germ. and preface by V. M. Ivanova. Moscow: Finansy i statistika, 1983, 304 p. [In Russian]

Дятлов Никита Евгеньевич

ассистент,
Медицинский институт,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: sakedas@gmail.com

Dyatlov Nikita Evgen'evich

assistant,
Medical Institute, Penza State University
(40 Krasnaya street, Penza, Russia)

Митрохина Наталья Юрьевна

кандидат технических наук, доцент,
кафедра теоретической и прикладной
механики и графики,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: natenal@yandex.ru

Mitrokhina Natal'ya Yur'evna

candidate of technical sciences,
associate professor,
sub-department of theoretical
and applied mechanics and graphics,
Penza State University
(40 Krasnaya street, Penza, Russia)

Рахматуллов Фазим Касымович

доктор медицинских наук, профессор,
заведующий кафедрой
внутренних болезней,
Медицинский институт,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: pgu-vb2004@mail.ru

Rakhmatullov Fagim Kasymovich

doctor of medical sciences, professor,
head of the department of internal diseases,
Medical Institute,
Penza State University
(40 Krasnaya street, Penza, Russia)

Митрохин Максим Александрович

доктор технических наук,
заведующий кафедрой
вычислительной техники,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: vt@pnzgu.ru

Mitrokhin Maksim Aleksandrovich

doctor of technical sciences,
head of the department
of computer engineering,
Penza State University
(40 Krasnaya street, Penza, Russia)

Образец цитирования:

Дятлов, Н. Е. Математические модели прогнозирования течения фибрилляции предсердий при беременности / Н. Е. Дятлов, Н. Ю. Митрохина, Ф. К. Рахматуллов, М. А. Митрохин // Модели, системы, сети в экономике, технике, природе и обществе. – 2019. – № 1-2 (29-30). – С. 139–146.