



ВРАЧ

И ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ

№3 2023

MEDICAL DOCTOR AND IT



MEDICAL REPORT
02-08-38 - MALE

02 :43 080
586 :89 403
253 :684 :01
99 :RP_809

ISSN 1811-0193
9 1771811019000 >

Журнал включен в Перечень рецензируемых научных изданий ВАК по специальностям:

- 2.3.1. Системный анализ, управление и обработка информации (технические науки);
- 2.3.5. Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей (технические науки);
- 3.3.9. Медицинская информатика (биологические науки);
- 3.3.9. Медицинская информатика (медицинские науки).

Журнал индексируется в базе данных Russian Science Citation Index (RSCI) на платформе Web of Science.

The journal is included in the Russian Science Citation Index (RSCI) database on the Web of Science platform.

ГЛАВНЫЙ РЕДАКТОР

Карпов О.Э., академик РАН, д.м.н., проф., генеральный директор ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России, Москва, Россия

ПОЧЕТНЫЙ ГЛАВНЫЙ РЕДАКТОР

Стародубов В.И., академик РАН, д.м.н., проф., научный руководитель ФГБУ «ЦНИИОИЗ» Минздрава России, представитель России в Исполнительном Комитете ВОЗ, Москва, Россия

ЗАМЕСТИТЕЛИ ГЛАВНОГО РЕДАКТОРА

Зарубина Т.В., д.м.н., член-корреспондент РАН, проф., заведующая кафедрой медицинской кибернетики и информатики, ФГАОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России, Москва, Россия

Гусев А.В., к.т.н., член экспертного совета Минздрава по вопросам использования ИКТ, старший научный сотрудник ФГБУ «ЦНИИОИЗ» Минздрава России, директор по развитию компании «К-Скай», Петрозаводск, Россия

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Андриков Д.А., к.т.н., доцент Инженерной Академии ФГАОУ ВО РУДН, директор компании «Иммersed», Москва, Россия

Владимирский А.В., д.м.н., заместитель директора по научной работе ГБУЗ «НПКЦ ДиТ ДЗМ», Москва, Россия

Грибова В.В., член-корреспондент РАН, д.т.н., заместитель директора по научной работе ФГБУ «Институт автоматизации и процессов управления» Дальневосточного отделения РАН, Владивосток, Россия

Гулиев Я.И., к.т.н., директор Исследовательского центра медицинской информатики ИПС РАН им. А.К. Айламазяна, Ярославль, Россия

Зингерман Б.В., руководитель направления цифровой медицины ИНВИТРО, Москва, Россия

Карась С.И., д.м.н., специалист отдела координации научной и образовательной деятельности НИИ кардиологии ФГБНУ «Томский национальный исследовательский медицинский центр Российской академии наук», профессор кафедры медицинской и биологической кибернетики ФГБОУ ВО СибГМУ Минздрава России, Томск, Россия

Лебедев Г.С., д.т.н., директор института цифровой медицины, заведующий кафедрой информационных и интернет технологий ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова» Минздрава России, Москва, Россия

Неусыпин К.А., д.т.н., профессор, заведующий кафедрой системы автоматического управления МГТУ им. Н.Э. Баумана, Москва, Россия

Пролетарский А.В., д.т.н., профессор, декан факультета «Информатика и системы управления» МГТУ им. Н.Э. Баумана, Москва, Россия

Реброва О.Ю., д.м.н., профессор кафедры медицинской кибернетики и информатики, ФГАОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России, Москва, Россия

Столбов А.П., д.т.н., профессор кафедры организации здравоохранения, медицинской статистики и информатики ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова» Минздрава России, Москва, Россия

Храмов А.Е., д.ф.м.н., профессор, руководитель Балтийского центра нейротехнологий и искусственного интеллекта Балтийского федерального университета им. И. Канта, Калининград, Россия

Швырев С.Л., к.м.н. заместитель руководителя Регламентной службы федерального реестра НСИ ФГБУ ЦНИИОИЗ Минздрава России, Москва, Россия

ИНОСТРАННЫЕ ЧЛЕНЫ РЕДАКЦИОННОЙ КОЛЛЕГИИ:

Писарчик А., к.б.н., проф., заведующий кафедрой вычислительной биологии, центр биомедицинских технологий, Мадридский технический университет, Мадрид, Испания

CHIEF EDITOR

Karpov O.E., Academician of the RAS, DSc, Prof., General Director of the Pirogov National Medical and Surgical Center, Moscow, Russia

HONORARY CHIEF EDITOR

Starodubov V.I., Academician of the RAS, DSc, Prof., Scientific Director of the FRIHOI of MoH of Russia, Representative of Russia in the WHO Executive Committee, Moscow, Russia

DEPUTY CHIEF EDITORS

Zarubina T.V., DSc, Corresponding Member of the RAS, Ptof., Head of the Department of Medical Cybernetics and Informatics, Pirogov Russian National Research Medical University, Moscow, Russia

Gusev A.V., PhD, member of the expert council of the Ministry of Health on the use of ICT, Senior Researcher of the FRIHOI of MoH of Russia, development director of the K-Sky company, Petrozavodsk, Russia

EDITORIAL BOARD

Andrikov D.A., PhD, Associate Prof. of the Engineering Academy of the RUDN University, Director of Immersed, Moscow, Russia

Vladimirsky A.V., DSc, Deputy Director for Research, Scientific and Practical Clinical Center for Diagnostics and Telemedicine Technologies Department of Health of the City of Moscow, Moscow, Russia

Gribova V.V., Corresponding Member of the RAS, DSc, Deputy Director for Research of the Federal State Budgetary Institution "Institute of Automation and Control Processes" of the Far Eastern Institute of the RAS Branch, Vladivostok, Russia

Guliev Ya.I., PhD, Director of the Research Center for Medical Informatics of the Institute of Applied Problems of the Russian Academy of Sciences named after A.K. Ailamazyan, Yaroslavl, Russia

Zingerman B.V., Head of Digital Medicine, INVITRO, Moscow, Russia

Karas S.I., DSc, Specialist of the Department for Research and Training Coordination, Cardiology Research Institute, Tomsk National Research Medical Centre, Russian Academy of Sciences; Professor at the Medical and Biological Cybernetics Chair, Siberian State Medical University, Tomsk, Russia

Lebedev G.S. DSc, Director of The Digital Health Institute, Head of The Department of information and internet technologies, I.M. Sechenov First Moscow State Medical University, Moscow, Russia

Neusypin K.A., DSc, Prof., Head of the Automatic Control Systems Dept., Bauman Moscow State Technical University, Moscow, Russia

Proletarsky A.V., DSc, Prof., Dean of the Informatics, and Control Systems Department, Bauman University, Moscow, Russia

Rebrova O.Yu., DSc, Prof. of the Department of Medical Cybernetics and Informatics, Pirogov Russian National Research Medical University, Moscow, Russia

Stolbov A.P., DSc, Prof. of the Department of Public Health Organization, Medical Statistics and Informatics of the Faculty of Professional Development of Doctors of the I.M. Sechenov First Moscow State Medical University, Moscow, Russia

Khramov A.E., DSc, Prof., Head of Baltic Center for Neurotechnology and Artificial Intelligence, Immanuel Kant Baltic Federal University, Kaliningrad, Russia

Shvyrev S.L., PhD, Deputy Head of the Regulatory Service of the Federal Register of the FRIHOI of MoH of Russia, Moscow, Russia

FOREIGN MEMBERS OF THE EDITORIAL BOARD:

Pisarchik A., PhD, Prof., Head of Department of Computational Biology, Center of Biomedical Technologies, Technical University of Madrid, Spain

Издается с 2004 года.

Включен в перечень ВАК ведущих рецензируемых научных журналов и изданий, рекомендуемых для опубликования основных научных результатов диссертации на соискание ученой степени кандидата и доктора наук.

Читатели могут принять участие в обсуждении статей, опубликованных в журнале «Врач и информационные технологии», и направить актуальные вопросы в редакцию (vit-j@pirogov-center.ru).

Журнал зарегистрирован Министерством Российской Федерации по делам печати, телерадиовещания и средств массовых коммуникаций.

Товарный знак и название «Врач и информационные технологии» являются исключительной собственностью ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России.

Авторы опубликованных материалов несут ответственность за подбор и точность приведенных фактов, цитат, статистических данных и прочих сведений, а также за то, что в материалах не содержится данных, не подлежащих открытой публикации.

Материалы рецензируются редакционной коллегией. Мнение редакции может не совпадать с мнением автора.

Перепечатка текстов без разрешения журнала «Врач и информационные технологии» запрещена. При цитировании материалов ссылка на журнал обязательна.

За содержание рекламы ответственность несет рекламодатель.

Учредитель — ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России.
Издатель — ФГБУ «НМХЦ им. Н.И. Пирогова» Минздрава России.

Адрес редакции:

105203, г. Москва,
ул. Нижняя Первомайская, д. 70,
e-mail: vit-j@pirogov-center.ru.
Тел. +7 (499) 464-03-03.

Главный редактор:

Карпов О.Э., академик РАН,
д.м.н., проф.

Почетный главный редактор:

Стародубов В.И.,
академик РАН, д.м.н., проф.

Зам. главного редактора:

Зарубина Т.В., член-корреспондент РАН,
д.м.н., проф.

Гусев А.В., к.т.н.

Компьютерная верстка и дизайн:

Издательство Пироговского Центра.

Подписные индексы:

Каталог агентства «Роспечать» — 82615.

Отпечатано в типографии ООО «Вива-Стар»
г. Москва, ул. Электrozаводская, д. 20
www.vivastar.ru

Подписано в печать 27 сентября 2023 г.
Общий тираж 1000 экз.

Распространяется бесплатно.

© Издательство Пироговского Центра

ОРИГИНАЛЬНЫЕ ИССЛЕДОВАНИЯ

*Бобровская Т.М., Кирпичев Ю.С., Савкина Е.Ф.,
Четвериков С.Ф., Арзамасов К.М.*

**РАЗРАБОТКА И ВАЛИДАЦИЯ ИНСТРУМЕНТА
СТАТИСТИЧЕСКОГО СРАВНЕНИЯ ХАРАКТЕРИСТИЧЕСКИХ
КРИВЫХ НА ПРИМЕРЕ РАБОТЫ АЛГОРИТМОВ НА ОСНОВЕ
ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА..... 4**

*Котловский М.Ю., Цыбикова Э.Б., Лорсанов С.М.,
Фадеев П.А., Фадеева С.О., Гусев А.В.*

**РАЗРАБОТКА МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ ДЛЯ
ПРОГНОЗИРОВАНИЯ ЧИСЛА ВПЕРВЫЕ ВЫЯВЛЕННЫХ
ПАЦИЕНТОВ С ВИЧ ИНФЕКЦИЕЙ В СУБЪЕКТАХ
РОССИЙСКОЙ ФЕДЕРАЦИИ 16**

Астанин П.А., Ронжин Л.В., Раузина С.Е.

**АЛГОРИТМ ОЦЕНКИ СПЕЦИФИЧНОСТИ ТЕРМИНОВ
МЕТАТЕЗАУРУСА UMLS НА ПРИМЕРЕ АНАЛИЗА
СЕМАНТИЧЕСКОЙ МОДЕЛИ ДЛЯ ДИФФЕРЕНЦИАЛЬНОЙ
ДИАГНОСТИКИ АКСИАЛЬНОГО СПОНДИЛОАРТРИТА..... 30**

Шахгельдян К.И., Гельцер Б.И., Куксин Н.С., Домжалов И.Г.

**МНОГОУРОВНЕВАЯ КАТЕГОРИЗАЦИЯ НЕПРЕРЫВНЫХ
ПЕРЕМЕННЫХ В ЗАДАЧАХ ОБЪЯСНЕНИЯ ПРОГНОЗНЫХ
ОЦЕНОК МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ В
КЛИНИЧЕСКОЙ МЕДИЦИНЕ..... 44**

ПРАКТИЧЕСКИЙ ОПЫТ

Журавлев Д.М., Копылов Ф.Ю., Чаадаев В.К., Ардатов С.В., Чаадаев К.В.

**АВТОМАТИЗИРОВАННЫЙ КОМПЛЕКС
МУЛЬТИДИСЦИПЛИНАРНОЙ НЕЙРОСЕТЕВОЙ ПОДДЕРЖКИ
ВРАЧЕБНЫХ РЕШЕНИЙ ПРИ ЛЕЧЕНИИ ИШЕМИЧЕСКОЙ
БОЛЕЗНИ СЕРДЦА..... 58**

*Климко В.И., Соломатников И.А., Ходакова О.В., Кунгурцев О.В.,
Страдымов Ф.И., Черномырдина Е.В.*

**ИСПОЛЬЗОВАНИЕ КАРТОГРАФИЧЕСКИХ СХЕМ
ТЕРРИТОРИАЛЬНОГО ПЛАНИРОВАНИЯ
В ЗДРАВООХРАНЕНИИ 72**

*Порецкова Г.Ю., Тяжева А.А., Плахотникова С.В., Чаплыгин С.С.,
Напалкова С.А., Баринов И.В., Панова Н.И.*

**ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ ЦИФРОВЫХ
ТЕХНОЛОГИЙ ПРИ ПРОФИЛАКТИЧЕСКОМ ОСМОТРЕ
ШКОЛЬНИКОВ Г. САМАРЫ 84**

ORIGINAL RESEARCH

*Bobrovskaya T.M., Kirpichev Y.S., Savkina E.F.,
Chetverikov S.F., Arzamasov K.M.*

**DEVELOPMENT AND VALIDATION OF A TOOL FOR STATISTICAL
COMPARISON OF ROC-CURVES USING THE EXAMPLE
OF ALGORITHMS BASED ON ARTIFICIAL INTELLIGENCE
TECHNOLOGIES..... 4**

*Kotlovskiy M.Yu., Tsybikova E.B., Lorsanov S.M.,
Fadeev P.A., Fadeeva S.O., Gusev A.V.*

**DEVELOPMENT OF A MACHINE LEARNING MODEL
PREDICTING THE INCIDENCE OF NEWLY DIAGNOSED
HIV INFECTION IN THE SUBJECTS
OF THE RUSSIAN FEDERATION 16**

Astanin P.A., Ronzhin L.V., Rauzina S.E.

**ALGORITHM FOR UMLS METATHESAURUS CONCEPTS
SPECIFICITY ESTIMATION USING EXAMPLE OF ANALYSIS
OF THE SEMANTIC MODEL DESCRIBING AXIAL
SPONDYLOARTHRITIS DIFFERENTIAL DIAGNOSTICS..... 30**

Shakhgeldyan K.I., Geltser B.I., Kuksin N.S., Domzhalov I.G.

**MULTILEVEL CATEGORIZATION OF CONTINUOUS
VARIABLES IN THE TASKS OF EXPLAINING PREDICTIVE
ESTIMATES OF MACHINE LEARNING MODELS
IN CLINICAL MEDICINE..... 44**

PRACTICE EXPERIENCE

Zhuravlev D.M., Kopylov F.Yu., Chaadaev V.K., Ardatov S.V., Chaadaev K.V.

**AUTOMATED COMPLEX OF MULTIDISCIPLINARY
NEURAL NETWORK SUPPORT OF MEDICAL
DECISION MAKING IN THE TREATMENT
OF CORONARY HEART DISEASE..... 58**

*Klimko V.I., Solomatnikov I.A., Khodakova O.V., Kungurtcev O.V.,
Stradymov F.I., Chernomyrdina E.V.*

**USING CARTOGRAPHIC SCHEME
OF TERRITORIAL PLANNING
IN HEALTH CARE..... 72**

*Poretskova G.Yu., Tyazheva A.A., Plakhotnikova S.V., Chaplygin S.S.,
Napalkova S.A., Barinov I.V., Panova N.I.*

**POSSIBLE APPLICATIONS OF DIGITAL TECHNOLOGIES
IN THE PREVENTIVE EXAMINATION OF SCHOOLCHILDREN
IN SAMARA 84**

MEDICAL DOCTOR AND INFORMATION TECHNOLOGIES

Registration certificate
PI No. FS77-80906 dated April 09, 2021

Published since 2004.

This journal is included in the list of the Higher Attestation Commission, detailing leading peer-reviewed scientific journals and publications recommended for publishing the foremost scientific results of dissertations for the degree of candidate and doctor of sciences.

Readers may take part in the discussion of articles published in the journal «Medical Doctor and Information Technologies», and send topical questions to the editorial office (vit-j@pirogov-center.ru).

The journal is registered by the Ministry of the Russian Federation for Press, TV and Radio Broadcasting, and Mass Media. The trademark and name «Medical Doctor and Information Technologies» are the exclusive property of the Pirogov National Medical and Surgical Center.

The authors of the published materials are responsible for the selection and accuracy of the facts, quotes, statistical data and other information, as well as ensuring that the materials do not contain data that is not subject to open publication.

The materials are reviewed by the editorial board. Editorial opinion may not reflect the views of the author.

Reprinting of texts without the permission of the journal «Medical Doctor and Information Technologies» is prohibited. When citing materials, a reference to the journal is required.

The advertiser is responsible for the content of the advertisement.

Founder — Pirogov National Medical and Surgical Center.

Publisher — Pirogov National Medical and Surgical Center.

Editorial office address:

105203, Moscow, st. Nizhnyaya Pervomayskaya, 70, e-mail: vit-j@pirogov-center.ru.
+7(499) 464-03-03.

Chief Editor:

Karpov O.E., Academician
of the Russian Academy of Sciences,
Doctor of Medical Sciences, Prof.

Honorary chief editor:

Starodubov V.I., Academician
of the Russian Academy of Sciences,
Doctor of Medical Sciences, Prof.

Deputy chief editors:

Zarubina T.V., Corresponding Member
of the Russian Academy of Sciences,
Doctor of Medical Sciences, prof.

Gusev A.V., Ph.D.

DTP and design:

Pirogov Center Publishing House.

Subscription indexes:

Catalogue of the agency
«Rospechat» — 82615.

Printed in the «Viva-Star»
Moscow, st. Elektrozavodskaya, 20
www.vivastar.ru

Signed for printing on September 27, 2023.

Circulation 1000 copies.

Free distribution.

© Pirogov Center Publishing House

БОБРОВСКАЯ Т.М.,

ГБУЗ «НПКЦ диагностики и телемедицины ДЗМ», Москва, Россия,
e-mail: BobrovskayaTM@zdrav.mos.ru

КИРПИЧЕВ Ю.С.,

ГБУЗ «НПКЦ диагностики и телемедицины ДЗМ», Москва, Россия,
e-mail: KirpichevYS@zdrav.mos.ru

САВКИНА Е.Ф.,

ГБУЗ «НПКЦ диагностики и телемедицины ДЗМ», Москва, Россия,
e-mail: SavkinaEF@zdrav.mos.ru, ORCID: 0000-0001-9165-0719

ЧЕТВЕРИКОВ С.Ф.,

к.т.н., ГБУЗ «НПКЦ диагностики и телемедицины ДЗМ», Москва, Россия,
e-mail: ChetverikovSF@zdrav.mos.ru

АРЗАМАСОВ К.М.,

к.м.н., ГБУЗ «НПКЦ диагностики и телемедицины ДЗМ», Москва, Россия; РТУ МИРЭА, Москва, Россия,
e-mail: ArzamasovKM@zdrav.mos.ru

РАЗРАБОТКА И ВАЛИДАЦИЯ ИНСТРУМЕНТА СТАТИСТИЧЕСКОГО СРАВНЕНИЯ ХАРАКТЕРИСТИЧЕСКИХ КРИВЫХ НА ПРИМЕРЕ РАБОТЫ АЛГОРИТМОВ НА ОСНОВЕ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

DOI: 10.25881/18110193_2023_3_4

Аннотация. Актуальность: Благодаря Национальной стратегии развития искусственного интеллекта в Российской Федерации активно разрабатываются и внедряются новые технологии на основе искусственного интеллекта, что приводит к появлению большого количества различных практических и научных задач, которые в свою очередь требуют удобных инструментов для их решения. Одним из них является инструмент, предназначенный для ROC-анализа, который был разработан и успешно применялся в рамках проекта «Эксперимент по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы». Однако для решения более широкого спектра задач, связанных с аналитикой работы технологий на основе искусственного интеллекта, возникла острая необходимость в разработке модуля сравнения ROC-кривых.

Цель: реализовать модуль инструмента ROC-анализа по сравнению площади под характеристической кривой с помощью статистических критериев и расчётом p -значения и апробировать его на реальных данных.

Материалы и методы: инструмент реализован на языке Python 3.9. 95% доверительный интервал для ROC-кривых рассчитывался с помощью метода случайных выборок с возвратом (бутстреппинг) и метода ДеЛонг (DeLong). Сравнение площадей под ROC-кривыми осуществлялось с помощью перестановочного теста.

Апробация инструмента осуществлялась на результатах работы 6 алгоритмов на основе технологий искусственного интеллекта на 2 наборах данных. Проводилось попарное сравнение площади под ROC-кривой и полученные результаты сравнивали с результатами анализа тех же данных методом ДеЛонг функции `roc.test` языка R 3.6.1.

Результаты: *p*-значения, полученные с помощью перестановочного теста, оказались в большинстве случаев сопоставимы с результатами *roc.test*, однако в 4 из 30 случаев *p*-значения принципиально отличались, что приводило к изменениям интерпретации теста.

Обсуждение: различия в результатах, рассчитанных двумя способами, вероятно, обусловлены особенностями используемых методов: ДеЛонг является более консервативным. Также из-за использования метода псевдорандомизации в перестановочном тесте возможна вариативность результатов, что может привести к неопределенности. Кроме того, разработанный инструмент сравнивает наборы данных с одинаковым количеством элементов, что является ограничением его использования, однако возможна дальнейшая его разработка с целью преодоления данного ограничения.

Заключение: был успешно реализован и апробирован модуль сравнения ROC-кривых с помощью статистических критериев с расчётом *p*-значения.

Ключевые слова: искусственный интеллект, ROC-анализ, статистический анализ.

Для цитирования: Бобровская Т.М., Кирпичев Ю.С., Савкина Е.Ф., Четвериков С.Ф., Арзамасов К.М. Разработка и валидация инструмента статистического сравнения характеристических кривых на примере работы алгоритмов на основе технологий искусственного интеллекта. *Врач и информационные технологии*. 2023; 3: 4-15. doi: 10.25881/18110193_2023_3_4.

BOBROVSKAYA T.M.,

State Budget-Funded Health Care Institution of the City of Moscow «Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department», Moscow, Russia, e-mail: BobrovskayaTM@zdrav.mos.ru

KIRPICHEV Y.S.,

State Budget-Funded Health Care Institution of the City of Moscow «Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department», Moscow, Russia, e-mail: KirpichevYS@zdrav.mos.ru

SAVKINA E.F.,

State Budget-Funded Health Care Institution of the City of Moscow «Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department», Moscow, Russia, e-mail: SavkinaEF@zdrav.mos.ru

CHETVERIKOV S.F.,

PhD, State Budget-Funded Health Care Institution of the City of Moscow «Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department», Moscow, Russia, e-mail: ChetverikovSF@zdrav.mos.ru

ARZAMASOV K.M.,

PhD, State Budget-Funded Health Care Institution of the City of Moscow «Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department», Moscow, Russia; Russian Technological University, Moscow, Russia, e-mail: ArzamasovKM@zdrav.mos.ru

DEVELOPMENT AND VALIDATION OF A TOOL FOR STATISTICAL COMPARISON OF ROC-CURVES USING THE EXAMPLE OF ALGORITHMS BASED ON ARTIFICIAL INTELLIGENCE TECHNOLOGIES

DOI: 10.25881/18110193_2023_3_4

Abstract. *Background: Due to the National Strategy for the Development of Artificial Intelligence, large-scale digitalization of healthcare is taking place in the Russian Federation, which leads to huge number of various practical and scientific tasks emergence of, which in turn require convenient tools to solve them. ROC analysis tool is one of them, which was developed and successfully applied within the framework of the project «Experiment on the use of innovative technologies in the field of computer vision for the analysis of medical images and further application in the healthcare system of the city of Moscow». However, there is an urgent need for the development of a module comparing ROC-curves in order to solve a wider range of problems related to analytics of the operation of technologies based on artificial intelligence.*

Aim: to implement the ROC analysis tool module for comparing the area under the curve using statistical methods and calculating the p-value, and to test it on real data.

Materials and methods: the tool is implemented in Python 3.9. The 95% confidence interval for ROC curves was calculated using the bootstrapping and the DeLong method. Areas under the ROC curves comparison was carried out using a permutation test.

The testing of the tool was carried out on the 6 algorithms work results on 2 data sets. Area under the ROC curve pairwise comparison was carried out and the results were compared with the same data results analysis, calculated by the DeLong method (roc.test function, R language 3.6.1).

Results: the p-values obtained using the permutation test were in most cases comparable to the roc.test results, however, in 4 out of 30 cases, the p-values differed significantly, which led to changes in the test interpretation.

Discussion: the differences in the results calculated by two separate methods, in our opinion, are due to the peculiarities of the methods used: DeLong method is more conservative. Also, due to the use of the pseudorandomization method in the permutation test, variability of results is possible, which can lead to uncertainty. In addition, the developed tool compares data of the same length, which is a limitation of its use, but its further development is possible for data of different lengths.

Conclusion: the module for comparing ROC curves was successfully implemented and tested using statistical criteria with the calculation of the p-value.

Keywords: artificial intelligence, ROC analysis, statistical analysis.

For citation: Bobrovskaya T.M., Kirpichev Y.S., Savkina E.F., Chetverikov S.F., Arzamasov K.M. Development and validation of a tool for statistical comparison of ROC-curves using the example of algorithms based on artificial intelligence technologies. *Medical doctor and information technology*. 2023; 3: 4-15. doi: 10.25881/18110193_2023_3_4.

ОБОСНОВАНИЕ

Разработка и внедрение программного обеспечения на основе технологий искусственного интеллекта (ТИИ) в РФ — одна из главных задач Национальной стратегии развития искусственного интеллекта в Российской Федерации [1]. Масштабная цифровизация всех сфер нашей жизни, в том числе и здравоохранения, приводит к увеличению количества данных и развитию инструментов по работе с ними [2]. Использование ТИИ в медицинской диагностике позволяет оптимизировать работу и снизить нагрузку на врачей. Например, во время пандемии COVID-19, использование алгоритмов на основе ТИИ позволило снизить время обработки протокола заключения компьютерной томографии врачом-рентгенологом [3]. Также исследования показали возможность использования таких алгоритмов в качестве второго чтения, например, при скрининге рака молочной железы [4]. Одним из самых масштабных проектов по внедрению ТИИ в здравоохранение является Московский «Эксперимент по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы» (далее Эксперимент) — наглядный пример успешной реализации применения ТИИ в медицинской диагностике [5].

Однако алгоритмы на основе ТИИ далеки от совершенства и требуют постоянного контроля качества, доработки и проверки. Прежде чем допустить такие алгоритмы к применению в медицинских организациях, необходимо провести тщательный анализ работы самого алгоритма, калибровку и оценку качества. Также необходим своевременный мониторинг качества в процессе работы алгоритма, который позволяет оперативно принимать решение о необходимости его перенастройки или вывода его из практической деятельности [6]. Одним из наиболее важных параметров является оценка диагностической точности. Золотым стандартом в оценке диагностических тестов является анализ классификации с помощью ROC-кривых (Receiver Operating Characteristic, характеристическая кривая).

Для реализации задач Эксперимента, а также в других научных исследованиях, требовался удобный и простой инструмент для оценки метрик диагностической точности: чувствительности, специфичности и площади под характеристической кривой ROC AUC (Receiver Operating Characteristic Area Under Curve). Для этого был реализован удобный, простой в использовании инструмент с открытым доступом [7, 8]. Однако недостаточно просто измерить абсолютные величины метрик диагностической точности. Ряд задач требует провести сравнение этих метрик между собой в соответствии с принципами доказательной медицины. Например, в задачах сравнения работы алгоритмов на основе ТИИ между собой или с работой врачей-рентгенологов. Также влияние различных факторов внешней среды может существенно изменить анализируемые данные, что в дальнейшем может привести к ухудшению метрик диагностической точности алгоритмов, а это, в свою очередь, потребует их дообучения, калибровки и нового тестирования для допуска к работе. Так, появление новой коронавирусной инфекции COVID-19 в 2020 году затруднило диагностику злокачественных новообразований легких [9] и, как следствие, мы столкнулись с задачей выявления статистически значимых различий в работе алгоритмов на основе ТИИ, чтобы ответить на вопрос: «Можем ли мы использовать данный алгоритм в новых условиях или необходима его доработка?».

ЦЕЛЬ

Реализовать модуль инструмента ROC-анализа по сравнению площади под характеристической кривой с помощью статистических критериев и расчётом р-значения и апробировать его на реальных данных.

МАТЕРИАЛЫ И МЕТОДЫ

Апробация нового модуля проводилась на данных, полученных в ходе зарегистрированного ранее исследования (NCT04489992), одобренного локальным этическим комитетом, «Эксперимент по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы».

Инструмент статистического анализа и построения характеристических кривых.

Интерфейс и серверная часть была написана с помощью библиотеки Plotly [10]. Для построения характеристических кривых с 95% доверительными интервалами с помощью метода случайных выборок с возвратом (далее «бутстреппинг» от англ. «bootstrapping») была использована библиотека ROC-utils [11]. Для вычислений доверительного интервала методом ДеЛонг была использована библиотека [12].

Сравнение площадей под характеристической кривой осуществлялось с помощью перестановочного теста [13]. Проверялась нулевая гипотеза об отсутствии статистически значимых различий между двумя ROC AUC.

Валидация инструмента.

С целью апробации работы нового инструмента мы провели ретроспективное обсервационное когортное исследование. Инструмент валидировался на результатах работы 6 алгоритмов на основе ТИИ (далее — ИИ1, ИИ2, ИИ3, ИИ4, ИИ5, ИИ6), полученных в ходе Эксперимента. Критерием для включения этих алгоритмов было соответствие заявленным возможностям определения солидных легочных узлов размером более 6 мм. Критерий исключения — проанализировано менее 90% исследований в каждом наборе данных (НД).

Алгоритмы обрабатывали оба НД в 2021 году одновременно без каких-либо доработок и изменений версий. Разработчики алгоритмов заявляли, что изменения, характерные для COVID-19, не являются ограничением применения их решений для обнаружения рака легких. Все исследования в каждом НД были обработаны всеми алгоритмами. Далее сравнивались результаты их работы между собой на каждом НД в отдельности с помощью готовой функции `roc.test` языка R (`method = «delong»`) и с помощью нашего инструмента. Проверялась нулевая гипотеза об отсутствии статистически значимых различий между алгоритмами на основе ТИИ.

Наборы данных.

Для исследования использовались следующие НД.

НД1 представляет собой 82 исследования компьютерной томографии грудной полости с наличием и отсутствием признаков рака легких (бинарная разметка, соотношение классов есть признаки патологии/нет признаков патологии 51/31), проведенных в период с 2015 по 2016, т.е. до начала пандемии COVID-19.

НД2 — 91 исследование компьютерной томографии грудной полости с наличием и отсутствием признаков рака легких (бинарная разметка, соотношение классов есть признаки патологии/нет признаков патологии 47/44), проведенных в 2020 году, в разгар пандемии. Соотношение признаков рака легкого и COVID-19: из 44 исследований без признаков рака легкого 22 исследования с признаками COVID-19, из 47 исследований с признаками рака легкого 29 исследований с признаками COVID-19.

Верификация НД проводилась путем экспертного пересмотра (исследования независимо анализировали 2 врача-рентгенолога, в случае разногласий подключался эксперт с опытом работы более 5 лет по данному направлению), а также с помощью патоморфологического исследования.

Критерии отнесения к классам с признаками/без признаков рака легких при экспертном пересмотре:

- с патологией: хотя бы один солидный или субсолидный узел объемом более 100 куб. мм; если волюметрию выполнить невозможно, использовался наименьший линейный размер узла 6 мм.
- без патологии: нет ни одного узла, подпадающего под указанные условия.

Верификация патологии COVID-19 для НД2 проводилась с помощью экспертного пересмотра, а также с помощью результатов лабораторного тестирования (ПЦР).

Критерии отнесения к классам с признаками/без признаков COVID-19 при экспертном пересмотре:

- с патологией:
 1. Инфильтрация легочной паренхимы по типу матовых стекол с обеих сторон, преимущественно периферической локализации, с или без инфильтрации легочной паренхимы по типу консолидации с положительным признаком воздушной бронхограммы;

2. Инфильтрация легочной паренхимы по типу булыжной мостовой (утолщение междолькового интерстиция на фоне матового стекла) с обеих сторон, преимущественно периферической локализации, с или без инфильтрации легочной паренхимы по типу консолидации с положительным признаком воздушной бронхограммы.

- без патологии: отсутствие вышеперечисленных признаков.

Критериями исключения для обоих НД являлись хирургические вмешательства, артефакты, связанные с пациентом (наложение руки на грудную клетку, ориентация тела, кашель, движения) или некачественное сканирование (артефакты, технические дефекты).

РЕЗУЛЬТАТЫ

Инструмент сравнения характеристических кривых.

Инструмент попарного сравнения результатов работы алгоритмов на основе ТИИ представляет собой специально созданный внутренний веб-инструмент (веб-сайт), на который необходимо загрузить файл с разметкой (уникальный идентификатор и соответствующее ему значение 1 или 0) и результатами, полученными от алгоритма.

После загрузки файлов отображаются их названия и результат сравнения ROC AUC методом

перестановочного теста (разница между значениями ROC AUC и р-значение для теста), а также непосредственно сами характеристические кривые. Кроме этого, инструмент оснащен функцией настройки параметров графика (цвет, оптимальное пороговое значение, отображение экспериментальных точек и доверительного интервала). Для характеристических кривых также рассчитывается и отображается площадь под кривой с доверительными интервалами, рассчитанными разными способами. Инструмент находится в открытом доступе по ссылке <https://roc-analysis.mosmed.ai>.

Результаты сравнения.

Для апробации созданного инструмента мы поставили задачу сравнить работу алгоритмов на основе ТИИ между собой в период до пандемии COVID-19 (НД1) и во время пандемии (НД2). Результаты попарного сравнения ROC AUC алгоритмов на основе ТИИ для выявления новообразований легких представлены в таблице 1. У 8 пар алгоритмов были обнаружены статистически значимые различия до пандемии и у 4 пар алгоритмов — во время пандемии.

Для сравнения работы инструментов расчета статистической значимости мы использовали р-значения, определенные с помощью функции `roc.test` языка R и перестановочного теста (Таблицы 2 и 3).

Таблица 1 — Результаты попарного сравнения ROC AUC алгоритмов на основе технологий искусственного интеллекта (ИИ1 — ИИ6). На главной диагонали — значения ROC AUC каждого алгоритма (желтый цвет). В ячейках указаны значения разности ROC AUC для каждого алгоритма (ИИi-ИИj). Зеленым цветом обозначены статистически значимые различия, определенные с помощью функции `roc.test` языка R. Таблица симметрична относительно главной диагонали (со сменой знака). А — результаты сравнения на НД1. Б — результаты сравнения на НД2

i \ j	ИИ1	ИИ2	ИИ3	ИИ4	ИИ5	ИИ6
ИИ1	0,761	-0,058	-0,166	-0,176	-0,181	-0,220
ИИ2	0,058	0,820	-0,107	-0,117	-0,122	-0,161
ИИ3	0,166	0,107	0,927	-0,010	-0,015	-0,054
ИИ4	0,176	0,117	0,010	0,937	-0,005	-0,044
ИИ5	0,181	0,122	0,015	0,005	0,942	-0,039
ИИ6	0,220	0,161	0,054	0,044	0,039	0,981

А

i \ j	ИИ1	ИИ2	ИИ3	ИИ4	ИИ5	ИИ6
ИИ1	0,897	0,133	0,122	0,086	0,093	0,013
ИИ2	-0,133	0,764	-0,011	-0,047	-0,040	-0,120
ИИ3	-0,122	0,011	0,775	-0,036	-0,029	-0,109
ИИ4	-0,086	0,047	0,036	0,811	0,007	-0,073
ИИ5	-0,093	0,040	0,029	-0,007	0,804	-0,080
ИИ6	-0,013	0,120	0,109	0,073	0,080	0,884

Б

Таблица 2 — р-значения для сравнения ROC AUC алгоритмов на основе ТИИ, рассчитанные с помощью функции roc.test языка R (таблица симметрична относительно главной диагонали). Красным цветом обозначены значения р, отличия в которых (по сравнению с перестановочным тестом) влияют на интерпретацию результата. А — результаты сравнения на НД1. Б — результаты сравнения на НД2

i \ j	ИИ1	ИИ2	ИИ3	ИИ4	ИИ5	ИИ6
ИИ1		0,409	0,014	0,005	0,001	0,000
ИИ2	0,409		0,038	0,024	0,017	0,001
ИИ3	0,014	0,038		0,744	0,718	0,080
ИИ4	0,005	0,024	0,744		0,888	0,054
ИИ5	0,001	0,017	0,718	0,888		0,204
ИИ6	0,000	0,001	0,080	0,054	0,204	

А

i \ j	ИИ1	ИИ2	ИИ3	ИИ4	ИИ5	ИИ6
ИИ1		0,007	0,021	0,107	0,054	0,762
ИИ2	0,007		0,844	0,396	0,291	0,025
ИИ3	0,021	0,844		0,510	0,605	0,008
ИИ4	0,107	0,396	0,510		0,898	0,139
ИИ5	0,054	0,291	0,605	0,898		0,145
ИИ6	0,762	0,025	0,008	0,139	0,145	

Б

Таблица 3 — р-значения для сравнения ROC AUC алгоритмов на основе технологий искусственного интеллекта, рассчитанные с помощью функции перестановочного теста (таблица симметрична относительно главной диагонали). Красным цветом обозначены значения р, отличия в которых (по сравнению с roc.test) влияют на интерпретацию результата. А — результаты сравнения на НД1. Б — результаты сравнения на НД2

i \ j	ИИ1	ИИ2	ИИ3	ИИ4	ИИ5	ИИ6
ИИ1		0,483	0,002	0,029	0,001	0,000
ИИ2	0,483		0,053	0,252	0,026	0,003
ИИ3	0,002	0,053		0,824	0,738	0,171
ИИ4	0,029	0,252	0,824		0,938	0,372
ИИ5	0,001	0,026	0,738	0,938		0,165
ИИ6	0,000	0,003	0,171	0,372	0,165	

А

i \ j	ИИ1	ИИ2	ИИ3	ИИ4	ИИ5	ИИ6
ИИ1		0,008	0,070	0,176	0,031	0,792
ИИ2	0,008		0,834	0,586	0,415	0,023
ИИ3	0,070	0,834		0,568	0,624	0,038
ИИ4	0,176	0,586	0,568		0,924	0,358
ИИ5	0,031	0,415	0,624	0,924		0,151
ИИ6	0,792	0,023	0,038	0,358	0,151	

Б

ОБСУЖДЕНИЕ

Одной из главных задач при разработке инструмента была задача выбора способа статистического сравнения площади под характеристической кривой. На сегодняшний день существует несколько способов статистического анализа метрик диагностической точности [14], а также различные инструменты их реализации, начиная от использования языков программирования и заканчивая готовыми

инструментами, не требующими знаний специальных языков [15–17]. Зачастую эти инструменты находятся в закрытом доступе, неудобны в использовании и сложны для понимания, кроме того, в случае самостоятельной реализации кода есть возможность его гибкой настройки под разные задачи. Поэтому мы поставили задачу реализовать удобный, доступный и простой в использовании инструмент. Ранее нами был реализован инструмент

построения ROC-кривых с расчётом доверительных интервалов для ROC AUC по методу ДеЛонга и с помощью бутстреппинга [8], и он успешно использовался в Эксперименте [18]. Однако в рамках задач по сравнению работы алгоритмов на основе ТИИ возникла необходимость в инструменте сравнения, который не только будет определять статистически обоснованные метрики, но и рассчитывать р-значение для сравниваемых метрик. Мы остановили свой выбор на перестановочном тесте [13]. Данный способ наиболее просто реализуется на языке Python, что было одним из важных условий при выборе языка программирования, т.к. наш инструмент изначально разрабатывался на нём. Еще одним аргументом в пользу метода перестановок является то, что он учитывает проблему множественных сравнений [19], что также играет важную роль, т.к. зачастую возникает необходимость сравнения большого количества моделей одновременно.

Для того чтобы оценить работу нашего инструмента, мы сравнивали полученные результаты с данными, рассчитанными с помощью функции `roc.test` языка R (Таблица 1,2), согласно которым у 8 пар алгоритмов на основе ТИИ имеются статистически значимые различия (р-значение меньше 0,05) до пандемии и у 4х пар алгоритмов — во время пандемии. Однако при расчете с помощью перестановочного теста мы получили несколько принципиальных расхождений, повлиявших на результаты интерпретации теста. Для НД1 р-значения пар ИИ2-ИИ3 и ИИ2-ИИ4 оказались выше 0,05, следовательно, мы не можем сделать вывод о наличии статистически значимой разницы между ними. Для НД2 р-значения пары ИИ1-ИИ3 оказались выше 0,05, а для ИИ1-ИИ5 — ниже, что также меняет интерпретацию результата на отсутствие статистически значимой разницы и ее наличие соответственно.

В целом, при анализе полученных с помощью перестановочного теста р-значений можно отметить их завышение по сравнению с `roc.test`, но в редких случаях значения совпадают или занижены. По мнению ряда авторов [20] метод ДеЛонга (использовался в функции `roc.test`) в ряде случаев является чрезмерно

консервативным и часто отвергает нулевую гипотезу об отсутствии статистически значимых различий в пользу альтернативной. Кроме того, результат метода ДеЛонга может зависеть от объема и баланса НД, на котором проводится исследование [20], что, вероятнее всего, и является причиной разнородного поведения р-значения при сравнении различных методов.

Стоит отметить тот факт, что перестановочный тест относится к категории симуляционных, поэтому он не обеспечивает полную воспроизводимость результата из-за псевдорандомизации (как и метод бутстреппинга, реализованный, например, в функции `roc.test`): выборка будет формироваться каждый раз случайным образом, поэтому р-значение будет колебаться в определенных пределах. Это может привести к ситуации неопределенности, когда р-значения будут больше или меньше 0,05 при повторении теста.

Таким образом, нам удалось реализовать модуль инструмента анализа характеристических кривых, позволяющий сравнить между собой площади под характеристическими кривыми, построенными на данных одной длины (количество исследований, которые анализировали алгоритмы на основе ТИИ). Возвращаясь к задаче, которая послужила нам примером для валидации инструмента, мы можем сказать, какие алгоритмы имеют статистически значимые различия. Однако нам бы хотелось найти ответ и на такой вопрос: изменились ли метрики качества работы алгоритмов на основе ТИИ во время пандемии COVID-19? И с помощью функции `roc.test` мы можем получить ответ, но в случае реализованного нами перестановочного теста мы не можем сравнивать НД с разным количеством элементов.

Разработанный нами инструмент может быть полезен для проведения клинических испытаний систем искусственного интеллекта [21], а также при оценке зрелости решений на основе ТИИ [22]. В дальнейшем мы планируем усовершенствовать наш инструмент и решить данную проблему, например, с помощью доэстраивания выборки по методу Гиббса [23], а также реализовать новые модули для решения различных задач анализа данных.

ЗАКЛЮЧЕНИЕ

Широкое использование алгоритмов на основе ТИИ приводит к появлению различных задач как для научных исследований, так и для контроля качества результатов их работы, что в свою очередь требует создания удобных инструментов анализа данных. Один из таких инструментов, предназначенный для построения и сравнения характеристических кривых, был реализован и успешно апробирован в данной работе. Однако эта реализация не работает для НД с разным количеством элементов, поэтому в дальнейшем планируется ее доработка.

Источник финансирования. Данная статья подготовлена авторским коллективом в рамках научно-исследовательской работы «Разработка платформы повышения качества ИИ-Сервисов для медицинской диагностики».

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Благодарности. Авторы выражают благодарность Никитину Н.Ю. за консультирование по вопросам статистического анализа.

ЛИТЕРАТУРА/REFERENCES

1. Указ Президента Российской Федерации от 10.10.2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации». 2019. Доступ по ссылке: <http://www.kremlin.ru/acts/bank/44731>. Ссылка активна на 14.02.2023. [Ukaz Prezidenta Rossijskoj Federacii ot 10.10.2019 g. «O razvitiu iskusstvennogo intellekta v Rossijskoj Federacii» № 490. Available at: <http://www.kremlin.ru/acts/bank/44731/page/1>. Accessed 14.02.2023. (In Russ.)]
2. Гусев А.В., Владимирский А.В., Шарова Д.Е. и др. Развитие исследований и разработок в сфере технологий искусственного интеллекта для здравоохранения в Российской Федерации: итоги 2021 года // Digital Diagnostics. — 2022. — Т.3. — №3. — С.178-194. [Gusev AV, Vladzimirskiy AV, Sharova DE, et al. Evolution of research and development in the field of artificial intelligence technologies for healthcare in the Russian Federation: results of 2021. Digital Diagnostics. 2022; 3(3): 178-194. (In Russ.)]. doi: 10.17816/DD107367.
3. Морозов С.П., Гаврилов А.В., Архипов И.В. и др. Влияние технологий искусственного интеллекта на длительность описаний результатов компьютерной томографии пациентов с COVID-19 в стационарном звене здравоохранения // Профилактическая медицина. 2022;25(1):14–20. [Morozov SP, Gavrillov AV, Arkhipov IV, et al. Effect of artificial intelligence technologies on the CT scan interpreting time in COVID-19 patients in inpatient setting. Profilakticheskaya Meditsina. 2022; 25(1): 14-20. (In Russ.)] doi: 10.17116/PROFMED20222501114.
4. Raya-Povedano JL, Romero-Martín S, Elías-Cabot E, et al. AI-based Strategies to Reduce Workload in Breast Cancer Screening with Mammography and Tomosynthesis: A Retrospective Evaluation. Radiology. 2021; 300(1): 57-65. doi: 10.1148/RADIOL.2021203555.
5. Морозов С.П., Владимирский А.В., Ледихова Н.В. и др. Московский эксперимент по применению компьютерного зрения в лучевой диагностике: вовлеченность врачей-рентгенологов // Врач и информационные технологии. 2020. — №4. — С.14-23. [Morozov SP, Vladzimirskiy AV, Ledikhova NV, et al. Moscow experiment on computer vision in radiology: involvement and participation of radiologists. Vrach i informacionnye tekhnologii. 2020; 4: 14-23. (In Russ.)]
6. Andreychenko AE, Logunova TA, Gombolevskiy VA, et al. A methodology for selection and quality control of the radiological computer vision deployment at the megalopolis scale. medRxiv. 2022: 2022.02.12.22270663. doi: 10.1101/2022.02.12.22270663.
7. Свидетельство о государственной регистрации программы для ЭВМ №2022617324 Российская Федерация. Веб-инструмент для выполнения ROC анализа результатов диагностических тестов: № 2022616046: заявл. 05.04.2022: опубл. 19.04.2022. С.П. Морозов, А.Е. Андрей-

- ченко, С.Ф. Четвериков и др. [Database registration certificate №2022617324 Web-instrument dlya vypolneniya ROC analiza rezul'tatov diagnosticheskikh testov: № 2022616046: Appl. 05.04.2022, publ. 19.04.2022. Morozov SP, Andreychenko AE, Chetverikov SF, et al. (In Russ.)]
8. ROC Analysis. Доступно по: <https://roc-analysis.mosmed.ai/> Ссылка активна на 12.08.2023. [ROC Analysis. Available at: <https://roc-analysis.mosmed.ai/> Accessed 12.08.2023. (In Russ.)]
 9. Goncalves S, Fong PC, Blokhina M. Artificial intelligence for early diagnosis of lung cancer through incidental nodule detection in low- and middle-income countries-acceleration during the COVID-19 pandemic but here to stay. *Am J Cancer Res.* 2022; 12(1): 1.
 10. Dash Documentation & User Guide Plotly. Available at: <https://dash.plotly.com/docs>. Accessed 08.08.2023.
 11. roc-utils. Available at: <https://github.com/hirsch-lab/roc-utils>. Accessed 21.08.2022.
 12. Sun X, Xu W. Fast implementation of DeLong's algorithm for comparing the areas under correlated receiver operating characteristic curves. *IEEE Signal Process Lett.* 2014; 21(11): 1389-1393. doi: 10.1109/LSP.2014.2337313.
 13. Pauly M, Asendorf T, Konietschke F. Permutation-based inference for the AUC: A unified approach for continuous and discontinuous data. *Biom J.* 2016; 58(6): 1319-1337. doi: 10.1002/BIMJ.201500105.
 14. Metz CE. ROC analysis in medical imaging: a tutorial review of the literature. *Radiol Phys Technol.* 2008; 1(1): 2-12. doi: 10.1007/S12194-007-0002-1/FIGURES/2.
 15. Statistical Software. Sample Size Software. NCSS. Available at: <https://www.ncss.com/> Accessed 08.02.2023.
 16. Goksuluk D, Korkmaz S, Zararsiz G, Karaagaoglu AE. EasyROC: An interactive web-tool for roc curve analysis using r language environment. *R Journal.* 2016; 8(2): 213-230. doi: 10.32614/RJ-2016-042.

17. ROC Analysis: Online ROC Curve Calculator. Available at: <http://www.rad.jhmi.edu/jeng/javarad/roc/JROCFITi.html>. Accessed 08.02.2023.
18. Artificial intelligence in radiology. Available at: <https://mosmed.ai/ai/> Accessed 08.02.2023.
19. Колядин В.Л. Пермутационные критерии как универсальный непараметрический подход к проверке статистических гипотез // Радиоэлектроника и информатика. — 2002. — №3. — С.20. [Kolyadin VL. Permutacionnyye kriterii kak universal'nyj neparametriceskij podhod k proverke statisticheskix gipotez. Radioelektronika i informatika. 2002; 3: 20. (In Russ.)]
20. Demler OV, Pencina MJ, D' RB, Sr A. Misuse of DeLong test to compare AUCs for nested models. Published online 2012. doi: 10.1002/sim.5328.
21. Клинические испытания систем искусственного интеллекта (лучевая диагностика) / сост. Ю.А. Васильев, А.В. Владзимирский, Д.Е. Шарова и др. // Серия «Лучшие практики лучевой и инструментальной диагностики». — Вып. 113. — 2-е изд., перераб. и доп. — М.: НПКЦ ДиТ ДЗМ, 2023. — 40 с. [Klinicheskie ispytaniya sistem iskusstvennogo intellekta (luchevaya diagnostika). Vasilyev YA, Vladzimirskyy AV, Sharova DE, et al. Seriya «Luchshie praktiki luchevoj i instrumental'noj diagnostiki». 2023. 40 p. (In Russ.)]
22. Тыров И.А., Васильев Ю.А., Арзамасов К.М и др. Оценка зрелости технологий искусственного интеллекта для здравоохранения: методология и ее применение на материалах московского эксперимента по компьютерному зрению в лучевой диагностике // Врач и информационные технологии. — 2022. — №4. — С.76-92. [Tyrov IA, Vasilyev YA, Arzamasov KM, et al. Assessment of the maturity of artificial intelligence technologies for healthcare: methodology and its application based on the use of innovative computer vision technologies for medical image analysis and subsequent applicability in the healthcare system of Moscow. Medical doctor and information technology. 2022; 4: 76-92. (In Russ.)] doi: 10.25881/18110193_2022_4_76. 21.
23. Probabilistic Graphical Models: Principles and Techniques — Daphne Koller, Nir Friedman.

КОТЛОВСКИЙ М.Ю.,

д.м.н., ФГБУ «ЦНИИОИЗ» Минздрава России, Москва, Россия,
e-mail: m.u.kotlovskiy@mail.ru

ЦЫБИКОВА Э.Б.,

д.м.н., ФГБУ «ЦНИИОИЗ» Минздрава России, Москва, Россия,
e-mail: erzheny2014@yandex.ru

ЛОРСАНОВ С.М.,

Министерство здравоохранения Чеченской Республики, г. Грозный, Россия,
e-mail: info@minzdravchr.ru

ФАДЕЕВ П.А.,

к.м.н., Министерство здравоохранения Чеченской Республики, г. Грозный, Россия,
e-mail: fadeipavel@mail.ru

ФАДЕЕВА С.О.,

Республиканский центр общественного здоровья и медицинской профилактики,
г. Грозный, Россия; Ярославский государственный медицинский университет,
г. Ярославль, Россия, e-mail: fadeeva-lana@inbox.ru

ГУСЕВ А.В.,

к.т.н., ФГБУ «ЦНИИОИЗ» Минздрава России, Москва, Россия, e-mail: agusev@webiomed.ai

РАЗРАБОТКА МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ЧИСЛА ВПЕРВЫЕ ВЫЯВЛЕННЫХ ПАЦИЕНТОВ С ВИЧ ИНФЕКЦИЕЙ В СУБЪЕКТАХ РОССИЙСКОЙ ФЕДЕРАЦИИ

DOI: 10.25881/18110193_2023_3_16

Аннотация. Цель: разработать модель прогнозирования числа впервые выявленных пациентов с ВИЧ-инфекцией в субъектах Российской Федерации с использованием методов машинного обучения

Материалы и методы: исходные данные были получены из формы федерального статистического наблюдения №61 и данных Росстата о среднегодовой численности населения - из 85 субъектов РФ (2016-2022 годы). Проведено сравнение методов машинного обучения и их ансамблей при построении регрессионной модели для прогнозирования числа впервые выявленных пациентов с ВИЧ-инфекцией в субъектах РФ.

Результаты: модель строилась с помощью методов: линейной регрессии, решающего дерева, случайного леса, градиентного бустинга на решающих деревьях и бэггинга. Использовалась интерактивная вычислительная среда «Jupyter Notebook» (6.5.2) и программные библиотеки «Pandas» (1.5.3), «Scikit-learn» (1.0.2), «Statsmodels» (0.13.5) и CatBoost. Оптимальные гиперпараметры подбирались с использованием фреймворка «Optuna». В качестве метрик качества выступили: корень из среднеквадратичной ошибки (RMSE); коэффициент детерминации (R²); средняя абсолютная ошибка (MAE); средняя абсолютная процентная ошибка (MAPE); медианная абсолютная ошибка (MedAE).

Выводы: применение методов и алгоритмов машинного обучения дает разные результаты в части метрик точности работы моделей. Наихудшие значения всех метрик качества продемонстрировал метод линейной регрессии (MAPE 67%). Наилучшим являлось сочетание (Бэггинг) двух ансамблевых методов — случайного леса и градиентного бустинга на решающих деревьях, поскольку было достигнуто максимальное значение большего числа метрик качества. В этой связи целесообразно проверять все доступные методы и алгоритмы машинного обучения и затем выбирать из полученных результатов наиболее качественную модель.

Ключевые слова: ВИЧ-инфекция, прогнозная аналитика, машинное обучение, искусственный интеллект.

Для цитирования: Котловский М.Ю., Цыбикова Э.Б., Лорсанов С.М., Фадеев П.А., Фадеева С.О., Гусев А.В. Разработка модели машинного обучения для прогнозирования числа впервые выявленных пациентов с ВИЧ инфекцией в субъектах Российской Федерации. *Врач и информационные технологии.* 2023; 3: 16-29. doi: 10.25881/18110193_2023_3_16.

KOTLOVSKIY M.YU.,

DSc, Central Research Institute of Organization and Informatization of Healthcare of the Ministry of Health of Russia, Moscow, Russia, e-mail: m.u.kotlovskiy@mail.ru

TSYBIKOVA E.B.,

DSc, Central Research Institute of Organization and Informatization of Healthcare» Ministry of Health of Russia, Moscow, Russia, e-mail: erzheny2014@yandex.ru

LORSANOV S.M.,

Ministry of Health of the Chechen Republic, Grozny, Russia, e-mail: info@minzdravchr.ru

FADEEV P.A.,

PhD, Ministry of Health of the Chechen Republic, Grozny, Russia, e-mail: fadeipavel@mail.ru

FADEEVA S.O.,

Republican Center for Public Health and Medical Prevention, Grozny, Russia; Yaroslavl State Medical University, Yaroslavl, Russia, e-mail: fadeeva-lana@inbox.ru

GUSEV A.V.,

PhD, Central Research Institute of Organization and Informatization of Healthcare of the Ministry of Health of Russia, Moscow, Russia, e-mail: agusev@webiomed.ai

DEVELOPMENT OF A MACHINE LEARNING MODEL PREDICTING THE INCIDENCE OF NEWLY DIAGNOSED HIV INFECTION IN THE SUBJECTS OF THE RUSSIAN FEDERATION

DOI: 10.25881/18110193_2023_3_16

Abstract. Aim: to develop a model predicting the incidence of newly diagnosed HIV infection in the subjects of the Russian Federation using machine learning methods.

Materials and methods: The initial data were obtained from the Federal statistical observation Form No. 61 and Rosstat data on the average annual population of 85 subjects of the Russian Federation (2016-2022). We made a comparison of machine learning methods and their ensembles in the construction of a regression model for predicting the incidence of newly diagnosed patients with HIV infection in the subjects of the Russian Federation.

Results: The model was built using the following methods: linear regression, decision Tree, random forest, gradient boosting on decision trees (GBDT) and bagging. The interactive computing environment «Jupyter Notebook» (6.5.2) and software libraries «Pandas» (1.5.3), «Scikit-learn» (1.0.2), «Statsmodels» (0.13.5) and CatBoost were utilized. Optimal hyperparameters were selected using the Optuna framework. The following quality metrics were used: root of mean square error (RMSE); coefficient of determination (R²); average absolute error (MAE); average absolute percentage error (MAPE); median absolute error (MedAE).

Conclusions: The use of machine learning methods and algorithms gives different results in terms of metrics of model accuracy. The worst values of all quality metrics were demonstrated by the linear regression method (MAPE 67%). The combination (bagging) of the two ensemble methods — Random Forest and GBDT — was the best, since the highest values were found for a larger number of quality metrics. In this regard, it is reasonable to test all available machine learning methods and algorithms and then select the best-quality model from the results obtained.

Keywords: *HIV infection, predictive analytics, machine learning, artificial intelligence.*

For citation: *Kotlovsky M.Yu., Tsybikova E.B., Lorsanov S.M., Fadeev P.A., Fadeeva S.O., Gusev A.V. Development of a machine learning model predicting the incidence of newly diagnosed HIV infection in the subjects of the Russian Federation. Medical doctor and information technology. 2023; 3: 16-29. doi: 10.25881/18110193_2023_3_16.*

ВВЕДЕНИЕ

Профилактика ВИЧ-инфекции среди населения подразделяется на первичную, направленную на исключение вероятности проникновения вирусных частиц в организм человека, и вторичную, при которой проводится профилактика и лечение заболеваний, способствующих заражению ВИЧ-инфекцией [1]. Одним из основных индикаторов, свидетельствующих об эффективности проводимых профилактических мер в субъектах Российской Федерации (РФ), является число впервые выявленных пациентов с ВИЧ-инфекцией [1]. Разработка методов моделирования для принятия управленческих решений, направленных на повышение эффективности проводимой первичной и вторичной профилактики ВИЧ-инфекции, является важной научно-практической задачей, в решении которой большую поддержку может оказать использование методов машинного обучения [2]. В настоящее время в ряде публикаций российских и зарубежных авторов представлены данные об использовании методов машинного обучения при построении прогнозных моделей для различных целевых событий, таких как оценка продолжительности жизни пациентов с ВИЧ-инфекцией, прогнозирование госпитальной летальности, оценка распространенности резистентности к антиретровирусным препаратам среди пациентов с ВИЧ-инфекцией и оценка риска заболеваемости ВИЧ-инфекцией [3–8]. Вместе с тем все еще остаются малоизученными вопросы, посвященные применению машинного обучения для создания моделей, позволяющих прогнозировать число впервые выявленных пациентов ВИЧ-инфекцией.

Цель исследования: разработать модель прогнозирования числа впервые выявленных пациентов с ВИЧ-инфекцией в субъектах Российской Федерации с использованием методов машинного обучения.

МАТЕРИАЛЫ И МЕТОДЫ

Для проведения исследования использованы данные из формы федерального статистического наблюдения №61 по 85 субъектам РФ за 2016–2022 годы и данные Росстата о среднегодовой численности населения данных субъектов РФ. Для построения прогностической модели были использованы следующие данные

1) среднегодовая численность населения субъектов РФ (Насел); 2) число лиц, обследованных с использованием методов лабораторной диагностики для выявления ВИЧ-инфекции в субъектах РФ (Блоттинг); 3) контингенты с ВИЧ-инфекцией, состоявшие под диспансерным наблюдением в СПИД-центрах в субъектах РФ на конец календарного года (Конт_ВИЧ).

Были рассчитаны следующие показатели:

1) Распространённость ВИЧ-инфекции

$$(\text{Распротр.} = \frac{\text{Конт. ВИЧ}}{\text{Насел.}} \times 100000);$$

2) Охват населения лабораторным обследованием для выявления ВИЧ-инфекции

$$(\text{ИФА_иссл.} = \frac{(\text{ИФА_иссл.})}{\text{Насел.}} \times 100000).$$

В качестве целевой переменной, используемой для прогноза при применении методов машинного обучения, явилось число впервые выявленных пациентов с ВИЧ-инфекцией в субъектах РФ.

Для работы с данными, полученными из 85 субъектов РФ, использовалась интерактивная вычислительная среда «Jupiter Notebook» (6.5.2) [9]. Обработка и анализ данных, а также составление и работа со структурированным датасетом производилась в программной библиотеке «Pandas» (1.5.3) [10]. Для построения линии тренда и нахождения коэффициентов линейной регрессии применялись программные библиотеки «Scikit-learn» (1.0.2) и «Statsmodels» (0.13.5) [11, 12].

Для построения прогностических моделей использованы методы линейной регрессии (Linear Regression), решающего дерева (Decision Tree), случайного леса (Random Forest), градиентного бустинга на решающих деревьях (GBDT), бэггинга и программные библиотеки Scikit-learn и CatBoost [13].

Оптимальные гиперпараметры прогностических моделей подбирались с использованием фреймворка «Optuna» [14].

В качестве метрик качества построенных моделей выступили: корень из среднеквадратичной ошибки (RMSE); коэффициент детерминации (R2); средняя абсолютная ошибка (MAE); средняя абсолютная процентная ошибка (MAPE); медианная абсолютная ошибка (MedAE). Для расчета

данных показателей использованы алгоритмы из модуля «Metrics» библиотеки «Scikit-learn» [15].

РЕЗУЛЬТАТЫ

В данном исследовании изначально было установлено, что значения **целевой объясняемой переменной** (число впервые выявленных пациентов с ВИЧ-инфекцией), значения которой мы прогнозировали, имели интервальную шкалу измерения. В связи с этим, методы машинного обучения, использованные для построения модели, относились к категории **регрессионных методов обучения с учителем**.

Целевая объясняемая переменная не имела нормального распределения. Это было подтверждено тестами Колмогорова-Смирнова и Шапиро-Уилка ($p < 0,05$). Распределение данной переменной по своей форме напоминало распределение Пуассона (рис. 1), что накладывало ограничение на применение ряда методов и могло снижать прогностическую точность построенных моделей. В связи с этим для улучшения качества прогноза целевая объясняемая переменная была преобразована путем нахождения натурального логарифма каждого из ее значений, что приближало ее распределение к нормальному. В свою очередь, после подбора оптимальных

гиперпараметров моделей, исходные параметры которых задает сам исследователь, и получения их предсказаний с использованием тестовых данных, для нахождения значений итоговых метрик качества проводилось обратное преобразование путем нахождения экспоненты каждого члена как преобразованных фактических данных, так и полученных предсказаний.

Первоначально была исследована предиктивная способность линейных и нелинейных одиночных методов машинного обучения и построенных на их основании моделей. Данными методами явились линейная регрессия и решающее дерево для регрессии [16–19].

Первая предиктивная модель была построена методом линейной регрессии, являющимся наиболее простым и изученным линейным методом машинного обучения, в котором предсказанные значения объясняемой целевой переменной y определяются через нахождение свободного члена b_0 , и коэффициентов $b_{1...k}$ для каждой из объясняющих переменных $X_{1...k}$, на основании которых делается прогноз в виде линии тренда. При этом математический алгоритм выстраивает прямую линию, максимально приближенную ко всем реальным значениям целевой объясняемой переменной $f(x, b) = b_0 + b_1 * X_1 + b_2 * X_2 + \dots + b_k * X_k$.

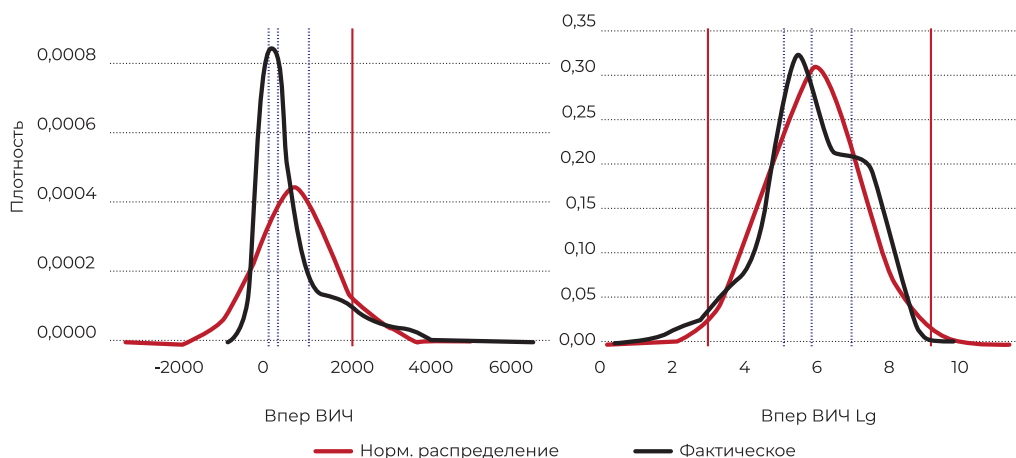


Рисунок 1 — Распределение значений целевой объясняемой переменной, 85 субъектов Российской Федерации, 2016–2022 годы, абсолютные значения. Здесь и далее: пунктирной вертикальной линией обозначены (квартили Q1, Q2, Q3) соответственно, сплошной вертикальной линией обозначены пределы доверительного интервала.

Данный метод является простым и интуитивно понятным. В нем отсутствует большое число гиперпараметров для настройки. Однако он демонстрирует хорошие результаты при наличии линейной зависимости между целевой и объясняющими переменными. При использовании данного метода для построения предиктивных моделей (но не для исследования связей) этот недостаток может быть преодолен за счет генерации дополнительных признаков путем проведения степенных преобразований значений

объясняющих переменных. Кроме того, к недостаткам данного метода относятся: повышенная чувствительность к разной масштабности показателей объясняющих переменных при нарушении их нормального распределения, наличие «выскакивающих» значений и сильно связанных между собой переменных [20].

В проведенном исследовании было установлено, что каждая из объясняющих переменных, на основе которых строился прогноз, как и объясняемая переменная, имели ненормальное

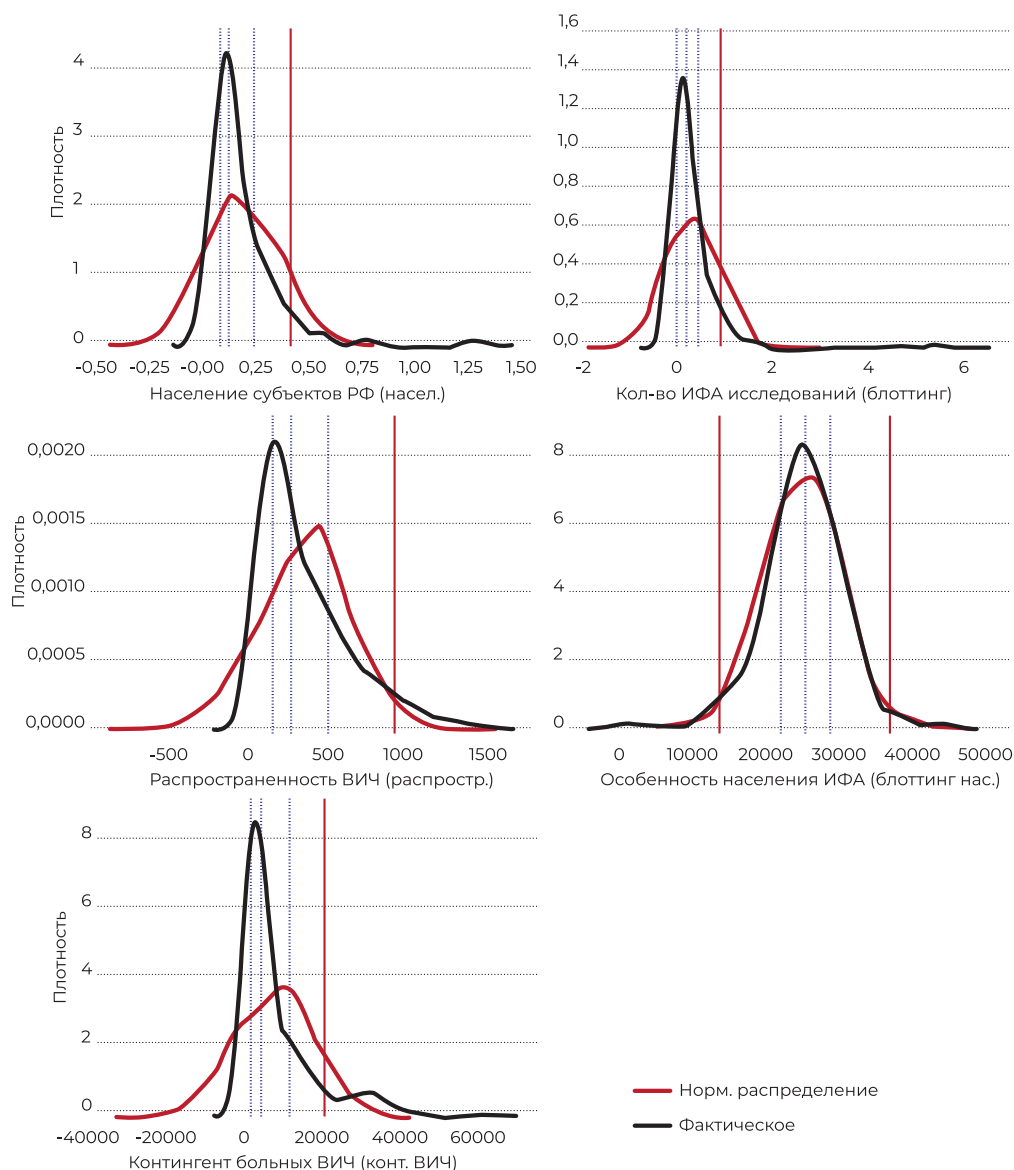


Рисунок 2 — Распределение значений объясняющих переменных, 85 субъектов Российской Федерации, 2016–2022 годы, абсолютные значения.

распределение из-за наличия выскакивающих значений (рис. 2). Это было подтверждено математически путем нахождения значений тестов Колмогорова-Смирнова и Шапиро-Уилка ($p < 0,05$). В дополнение к этому значения переменных лежали в разных числовых диапазонах.

В связи с вышеперечисленными особенностями используемых данных для применения метода линейной регрессии (и других методов), была проведена их подготовка. Для этого было произведено робастное шкалирование. Для этого использовали алгоритм `RobustScaler` из программного модуля `Preprocessing` бесплатной библиотеки машинного обучения для языка программирования Python-`Scikit-learn` [21]. При этом из каждого значения удалялась медиана (определенная на тренировочных данных), и происходило масштабирование значения в соответствии с интерквартильным размахом (тренировочных данных), диапазоном между 1-м квартилем (25-й квартиль) и 3-м квартилем (75-й квартиль).

Данное шкалирование применялось и в работе с остальными методами машинного обучения. Поскольку переменные закономерно сильно коррелировали между собой, применяли дополнительно L1 регуляризацию (`Lasso` — регрессию), в основе которой лежит идея добавления штрафного коэффициента к первоначальной функции потерь, что позволяет произвести разрежение и из всего массива объясняющих переменных отобрать наиболее важные, задающие тенденцию, удалив все остальное (шум) [22]. В результате из модели в автоматическом режиме были исключены коррелирующие переменные и переменные с низкой предсказательной ценностью. Для этого была использована `Lasso` модель из модуля — «`Linear_model`» библиотеки «`Scikit-learn`» [18].

Далее был проведен поиск оптимального значения единственного гиперпараметра — α . Это значение штрафного коэффициента для функции потерь, которое по умолчанию было равно 1. Подбор значений производился в диапазоне от 0,001 до 20. Для подбора оптимального значения гиперпараметров здесь и далее применяли фреймворк `Optuna`, предназначенный для автоматизированного поиска оптимальных значений гиперпараметров методов машинного обучения [14]. В качестве функции потерь была использована функция вычисления

среднеквадратичной ошибки (MSE), конкретно в данном случае с L1 регуляризацией. Предиктивная эффективность каждой модели проверялась методом кросс валидации на 5 фолдах (блоках) путем максимизации среднего значения негативной среднеквадратичной логарифмической ошибки (MSLE) [23]. Благодаря большому «гасящему» эффекту логарифма, данная функция потерь более применима к данным, имеющим разброс значений в несколько порядков, что имело место и в нашем случае. Для нахождения наилучшего значения штрафного коэффициента было проведено 10000 итераций (пробных построений модели). Это заняло 16 минут. Наилучшее найденное значение гиперпараметра α было установлено на 3839 итерации и составило 0,001, то есть наилучшие значения предсказаний были получены при штрафном коэффициенте, стремящемся к нулю, или полном отсутствии регуляризации.

Каждый анализируемый признак имел свою предсказательную ценность (рис. 3). Среди них основной ценностью обладали такие признаки, как численность населения (37%) и распространённость ВИЧ-инфекции (33%). Далее следовало число проведенных обследований (16%) и численность контингентов с ВИЧ-инфекцией (13%). Предсказательная ценность практически отсутствовала у признака — обеспеченность населения лабораторным обследованием для выявления ВИЧ-инфекции (1%).

В построенном уравнении регрессии свободный член был равен 5,52. Установленные коэффициенты были положительными для следующих признаков: численность населения — 1,51; обеспеченность населения лабораторным обследованием — 0,05; распространённость ВИЧ-инфекции — 1,37. В тоже время значения ряда коэффициентов были отрицательными для таких признаков, как число проведенных обследований -0,65 и численность контингентов с ВИЧ-инфекцией -0,54, что могло быть обусловлено наличием сильных корреляционных связей. Удаление одной из коррелирующих переменных снижало предсказательную ценность модели, как и построение новых полиномиальных и обобщенных моделей на основе имеющихся переменных.

После подбора оптимальных гиперпараметров модели выборку разделяли на обучающую

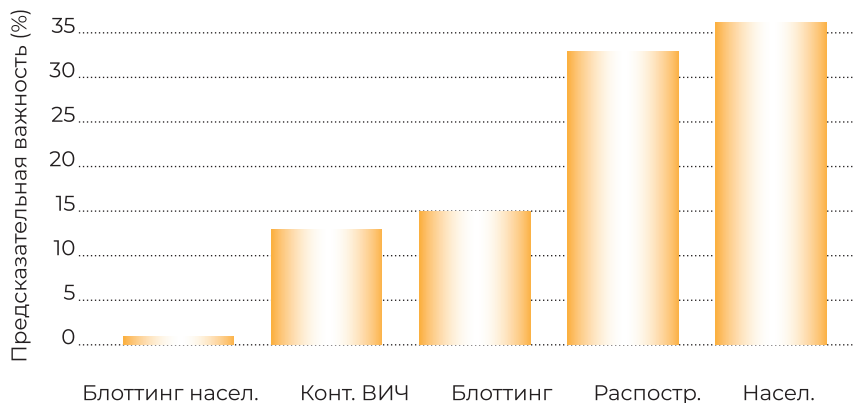


Рисунок 3 — Предсказательная ценность переменных при построении линейной регрессионной модели (с L1 регуляризацией).

и тестовую в соотношении 9 к 1. Производили обучение модели и делали прогноз. После этого вычисляли значение метрик качества, используя экспоненту предсказанных и фактических значений (операция обратная ранее проведенному логарифмированию).

Вторым примененным методом машинного обучения было решающее дерево [19]. Это алгоритм используется для построения одиночной и, в данном случае, нелинейной предиктивной модели. Устройство данного дерева включает в себя так называемые «ветви» и «листья». На «ветвях» решающего решения записаны признаки, от которых зависит значение целевой переменной, а в «листьях», которыми заканчиваются «ветви», записаны ее значения. Чтобы сделать прогноз, необходимо опуститься по дереву вплоть до листа и получить соответствующее значение. Метод отличается простота и высокая (не уступающая линейной регрессии) скорость построения модели, а также интуитивная понятность и наличие возможности графического отображения. Он менее требователен к наличию линейной связи, нормальности распределения, отсутствию выбросов, одинаковому масштабу данных и отсутствию корреляций.

Для данного метода использовалась базовая модель «DecisionTreeRegressor» из модуля «Tree» библиотеки «Scikit-learn» и производился подбор таких гиперпараметров как:

- глубина решающего дерева (max_depth) в диапазоне от 2 до 30;

- минимальное число наблюдений, необходимое для разделения внутреннего узла (min_samples_split), в диапазоне от 1 до 50;
- минимальное число наблюдений, необходимых для образования листа (min_samples_leaf), в диапазоне от 1 до 50 [19].

Всего было проведено 10000 итераций в течение 34 минут. Наилучшее сочетание гиперпараметров было установлено на 5056 итерации.

Несмотря на то, что данная модель не требовала нормального распределения и была устойчива к наличию выбросов, применяли масштабирование используя RobustScaler [21]. В результате подобранная модель с наилучшими показателями имела следующие характеристики: 1) max_depth — 7; 2) min_samples_split — 25; 3) min_samples_leaf — 5.

При этом каждая объясняющая переменная имела свою прогностическую ценность, основной из которых обладала такая переменная, как численность контингентов с ВИЧ-инфекцией (94%). Более низкую значимость показала переменная — численность населения (4%). Прогностическая ценность других переменных имела следовые значения (охват обследованием -0,9%, распространенность ВИЧ-инфекции -0,7%, обеспеченность населения лабораторным обследованием -0,4%).

Далее была изучена предиктивная способность моделей, построенных на основании более сложных нелинейных ансамблевых методов и их сочетаний. **Третья** предиктивная модель была построена на основании



Рисунок 4 — Предсказательная ценность переменных при построении модели «решающего дерева», 85 субъектов Российской Федерации, 2016–2022 годы, абсолютные значения.

алгоритма случайного леса [24]. При этом основной идеей является использование большого количества «решающих деревьев», каждое из которых в отдельности даёт относительно невысокое качество прогноза, но их общее число делает суммарный результат более высоким. Особенностью этого метода является смещение значения полученного прогноза при низком разбросе, вызванном переобучением. Кроме того, каждое «решающее дерево» достраивается независимо от другого, что даёт возможность распараллелить вычисления.

Для данного метода использовалась модель «RandomForestRegressor» из модуля «Ensemble» библиотеки «Scikit-learn» [25]. Данные подвергались робастному шкалированию. Производился подбор оптимальных значений следующих гиперпараметров:

- параметр сложности, используемый для снижения сложности с минимальными затратами (ccr_alpha), в диапазоне от 0,0001 до 5;
- максимальная глубина дерева (max_depth) от 2 до 30;
- количество деревьев в ансамбле ($n_estimators$) от 1 до 10000;
- минимальное число наблюдений, необходимое для разделения внутреннего узла ($min_samples_split$), от 2 до 50;
- минимальное число наблюдений, необходимых для образования листа ($min_samples_leaf$), от 1 до 30;

- минимальная величина примесей ($min_impurity_decrease$) от 0,001 до 5 ($log = True$).

Было проведено 10000 итераций в течение 2 часов 49 минут. Наилучшее сочетание гиперпараметров было установлено на 5056 итерации.

В результате подобранная модель с лучшими показателями имела следующие характеристики: 1) ccr_alpha — 0,0001; 2) max_depth — 8; 3) $n_estimators$ — 2112; 4) $min_samples_split$ — 3; 5) $min_samples_leaf$ -2; 6) $min_impurity_decrease$ 0,001.

Было установлено, что, как и в случае построения простого «решающего дерева», основной предсказательной ценностью обладала переменная — контингенты с ВИЧ-инфекцией (91%) (рис. 5). Низкую ценность показали переменные — численность населения (5%), число обследованных лиц (2%), распространённость ВИЧ-инфекции (1%) и обеспеченность населения лабораторным обследованием (1%).

На следующем шаге для построения предиктивной модели использован более сложный алгоритм машинного обучения, построенный на принципах градиентного бустинга над решающими деревьями [26]. Бустинг — это техника ансамблей методов машинного обучения, в основе которой лежит последовательное обучение нескольких моделей для повышения точности всей системы [13]. Этот метод использует идею о том, что каждая последующая модель будет учиться на ошибках предыдущей. Таким

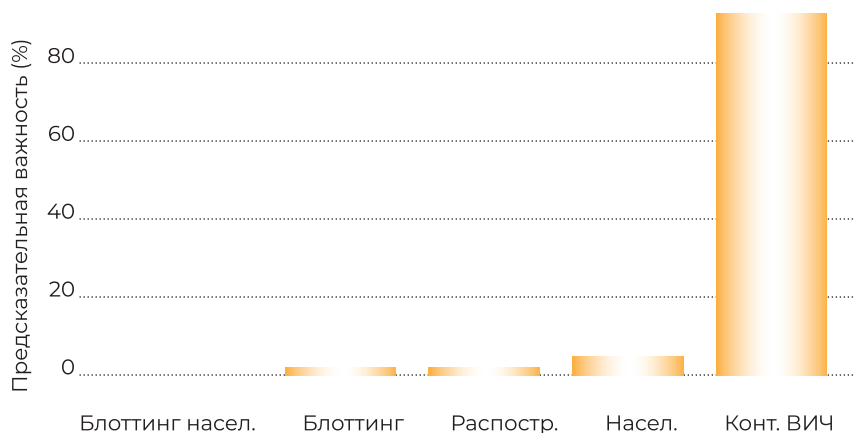


Рисунок 5 — Предсказательная важность переменных при построении предиктивной модели «случайный лес», 85 субъектов Российской Федерации, 2016–2022 годы, абсолютные значения.

образом, построение «решающих деревьев» в нашем случае происходит последовательно одно за другим, что требует гораздо больших временных затрат, чем предыдущие методы. Также особенностями данного метода, с одной стороны, является низкое смещение предсказанного значения, а с другой — склонность к переобучению и высокому значению разброса предсказаний.

Нами использована отечественная разновидность метода машинного обучения — «CatBoostRegressor» из программной библиотеки «CatBoost» [13]. Произведен подбор оптимальных значений следующих гиперпараметров модели:

- штрафного коэффициента ('penalties_coefficient) в диапазоне от 0.1 до 10;
- максимальной глубины дерева (max_depth) от 2 до 16;
- количества деревьев в ансамбле (num_trees) от 1 до 10000;
- минимального числа наблюдений, необходимых для образования листа (min_data_in_leaf), от 1 до 30.

Было проведено 100 итераций в течение 13 часов 30 мин. Наилучшие показатели были получены на 78 итерации.

В результате подобранная модель с наилучшими показателями имела следующие характеристики: 1) penalties_coefficient — 5,56;

2) max_depth — 5; 3) num_trees — 3224; 4) min_data_in_leaf — 6.

Было установлено, что наибольшей предсказательной ценностью обладала переменная — контингенты с ВИЧ-инфекцией (44%) (рис. 6).

Далее в порядке убывания следовали: численность населения (24%), число обследованных лиц (14%), распространённость ВИЧ-инфекции (11%) и обеспеченность населения лабораторным обследованием (7%). При этом полученные значения каждой из используемых метрик качества, за исключением медианной ошибки, превышали таковые у предшествующих моделей.

Следующим использованным методом стал Бэггинг, суть которого заключается в комбинации предсказанных значений независимых методов [27]. В работе был произведен расчет предсказательной способности обоих проанализированных ансамблевых методов — случайного леса и градиентного бустинга на решающих деревьях. При этом производили сложение предсказанных значений в подобранных оптимальных соотношениях $y = 0,3\text{RandomForest} + 0,7\text{GBDT}$.

ОБСУЖДЕНИЕ

Полученные значения метрик качества каждой из построенных прогностических моделей представлены в таблице 1. При этом мы

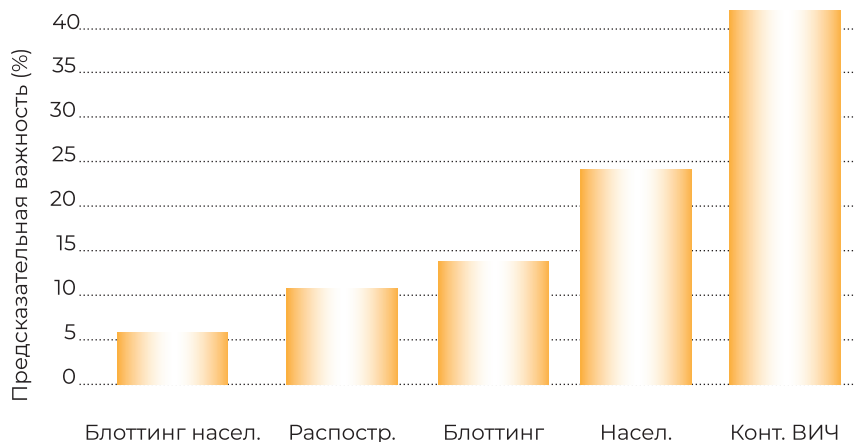


Рисунок 6 — Распределение переменных в зависимости от их прогностической ценности при построении предиктивной модели «градиентный бустинг на решающих деревьях», 85 субъектов Российской Федерации, 2016–2022 годы, абсолютные значения.

Таблица 1 — Метрики качества построенных предиктивных регрессионных моделей

Метрика качества	Линейная регрессия + (L1 – рег.)	Решающее дерево	Случайный лес	градиентный бустинг на решающих деревьях	Случайный лес + градиентный бустинг на решающих деревьях
R2, %	27,51	76,96	88,62	90,74	94,31
RMSE абс. число	827,46	466,57	327,85	295,69	293,91
MAE абс. число	349,92	274,77	213,25	190,25	192,48
MAPE, %	67	38	30,79	30,45	29,97
MedAE абс. число	119,18	99,78	90,67	103,34	114,71

старались получить наибольшее максимально приближенное к 100 % значение R2 и наименьшие значения остальных показателей, говорящих об абсолютном (RMSE, MAE, MedAE) или относительном (MAPE) значении величины ошибки.

Используемые метрики качества были подобраны на основании интуитивной понятности. При этом каждая из них могла иметь свои недостатки, поэтому прогностическая ценность построенных моделей оценивалась комплексно. Наихудшие значения всех метрик качества продемонстрировал метод линейной регрессии (MAPE 67%). Наилучшим было сочетание (Бэггинг) двух ансамблевых методов: случайного

леса, градиентного бустинга на решающих деревьях, поскольку было установлено наилучшее значение большего числа метрик качества (R², RMSE, MAPE). У ансамблевого метода случайный лес наилучшим среди остальных было значение метрики качества — MedAE. Хорошо проявила себя отечественная разновидность метода градиентного бустинга на решающих деревьях — алгоритм CatBoost, наилучшим было значение показателя MAE. При этом данный метод показал лучшие, чем случайный лес, значения четырех из пяти метрик качества (R², RMSE, MAPE, MAPE). Однако конструктивные особенности применяемого метода требовали значительно расхода времени для подбора оптимальных

гиперпараметров и обучения модели. Также требовалась высокая производительность ЭВМ. Это может стать преградой для включения дополнительных объясняющих переменных и увеличения числа наблюдений при построении модели. В тоже время применение виртуальных машин таких сервисов, как «Google Colab», «Kaggle», «Paperspace Gradient», «Deepnote», «Yandex DataSphere» и др., для повышения вычислительных возможностей может быть ограничено лишь характером анализируемых данных.

Выводы

Результаты проведенного исследования показали, что применение современных методов и алгоритмов машинного обучения может давать различные результаты в части метрик

точности работы моделей, поэтому при решении прикладных задач целесообразно проверять все доступные методы и алгоритмы машинного обучения и затем выбирать из полученных результатов наиболее качественную модель. Такой подход требует большие затраты времени на подготовку данных, проведение экспериментов с различными вариантами машинного обучения, а затем дополнительную настройку и поиск гиперпараметров в выбранной версии модели. В этой связи целесообразно проводить дальнейшие исследования в части применения технологий автоматизированного машинного обучения (AutoML), а также применения высокопроизводительных вычислительных комплексов и специализированных платформ, ускоряющих процессы обучения.

ЛИТЕРАТУРА/REFERENCES

1. ВИЧ-инфекция и СПИД. Национальное руководство. Под ред. акад. РАН, профессора В.В. Покровского, Москва: ГЭОТАР-МЕДИА, 2020. — 686 с. [HIV infection and AIDS. National leadership. Acad. RAS, Professor V.V. Pokrovsky, editor. Moscow: GEOTAR-MEDIA, 2020. 686 p. (In Russ.)]
2. Бодрин К.А., Красноперова А.А. Использование технологий машинного обучения в медицине // Теория и практика современной науки. — 2018. — №10(40). — С.52-56. [Bodrin KA, Krasnoperova AA. The use of machine learning technologies in medicine. Theory and practice of modern science. 2018; 10(40): 52-56. (In Russ.)]
3. Вострокнутов М.Е., Дюжева Е.В., Кузнецова А.В., Сенько О.В. Факторы риска госпитальной летальности больных с сочетанием туберкулеза и ВИЧ-инфекции в учреждениях уголовно-исполнительной системы // Туберкулез и болезни легких. — 2019. — Т.97. — №7. — С.34-41. [Vostroknutov ME, Dyuzheva EV, Kuznetsova AV, Senko OV. Risk factors of hospital mortality of patients with a combination of tuberculosis and HIV infection in institutions of the penal system. Tuberculosis and lung diseases. 2019; 97(7): 34-41. (In Russ.)] doi: 10.21292/2075-1230-2019-97-7-34-41.
4. Тарасова О.А., Филимонов Д.А., Поройков В.В. Компьютерный прогноз резистентности вируса иммунодефицита человека к ингибиторам обратной транскриптазы ВИЧ // Биомедицинская химия. — 2017. — Т.63. — №5. — С.457-460. [Tarasova OA, Filimonov DA, Poroikov VV. Computer prediction of human immunodeficiency virus resistance to HIV reverse transcriptase inhibitors. Biomedical chemistry. 2017. 63(5): 457-460. (In Russ.)] doi: 10.18097/PBMC20176305457.
5. Rajendran M, Ferran MC, Mouli L, Babbitt GA. Lynch Evolution of drug resistance drives destabilization of flap region dynamics in HIV-1 protease. *Biophys Rep (NY)*. 2023; 3(3): 100121. doi: 10.1016/j.bpr.2023.100121.
6. Bukic E, Milasin J, Toljic B, Jadzic J, Jevtovic D, Obradovic B, Dragovic G. Association between Combination Antiretroviral Therapy and Telomere Length in People Living with Human Immunodeficiency Virus. *Biology (Basel)*. 2023; 12(9): 1210. doi: 10.3390/biology12091210.
7. Birri Makota RB, Musenge E. Predicting HIV infection in the decade (2005-2015) pre-COVID-19 in Zimbabwe: A supervised classification-based machine learning approach. *PLOS Digit Health*. 2023; 2(6): e0000260. doi: 10.1371/journal.pdig.0000260.
8. Mamo DN, Yilma TM, Fekadie M, Sebastian Y, Bizuayehu T, Melaku MS, Walle AD. Machine learning to predict virological failure among HIV patients on antiretroviral therapy in the University of Gondar Comprehensive and Specialized Hospital, in Amhara Region, Ethiopia, 2022. *BMC Med Inform Decis Mak*. 2023; 23(1): 75. doi: 10.1186/s12911-023-02167-7.

9. Jupyter Notebook. Available at: <https://docs.jupyter.org/en/latest/>. Accessed 10.10.2023.
10. Pandas. Available at: <https://pandas.pydata.org/docs/>. Accessed 10.10.2023.
11. Scikit-learn. Documentation. Available at: <https://scikit-learn.org/stable/index.html>. Accessed 10.10.2023.
12. Statsmodels. Available at: <https://www.statsmodels.org/stable/user-guide.html>. Accessed 10.10.2023.
13. CatBoost. Available at: <https://catboost.ai/en/docs/>. Accessed 10.10.2023.
14. Optuna. Available at: https://optuna.org/#key_features. Accessed 10.10.2023.
15. Scikit-learn. Evaluation of models. Available at: https://scikit-learn.org/stable/modules/model_evaluation.html. Accessed 10.10.2023.
16. Лысенко А.А. Введение в регрессионный анализ данных и регрессионные модели // Труды Санкт-Петербургского государственного морского технического университета. — 2020. — Т.1. — №S2. — С.15. [Lysenko AA. Introduction to regression analysis of data and regression models. Proceedings of the St. Petersburg State Maritime Technical University. 2020; 1(S2): 15. (In Russ.)]
17. Пернебай Б.А. Python. Регрессия дерева решений с использованием sklearn // Polish Journal of Science. — 2021. — №38-1(38). — С.51-56. [Pernebai BA. Python. Decision tree regression using sklearn. Polish Journal of Science. 2021; 38-1(38): 51-56. (In Russ.)]
18. Scikit-learn. Linear models. Available at: https://scikit-learn.org/stable/modules/classes.html#module-sklearn.linear_model. Accessed 10.10.2023.
19. Scikit-learn. Decision tree, regressor. Available at: <https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeRegressor.html#sklearn.tree.DecisionTreeRegressor>. Accessed 10.10.2023.
20. Scikit-learn. Common errors in the interpretation of linear model coefficients. Available at: https://scikit-learn.org/stable/auto_examples/inspection/plot_linear_model_coefficient_interpretation.html#sphx-glr-auto-examples-inspection-plot-linear-model-coefficient-interpretation-py. Accessed 10.10.2023.
21. Scikit-learn. Robust scaling. Available at: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.RobustScaler.html>. Accessed 10.10.2023.
22. Scikit-learn. Lasso regression. Available at: scikit-learn.org/stable/modules/generated/sklearn.linear_model.Lasso.html. Accessed 10.10.2023.
23. Scikit-learn. Cross-validation. Available at: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.cross_validate.html. Accessed 10.10.2023.
24. Носова Г.С., Абдуллин А.Х. Машинное обучение на основе непараметрического и нелинейного алгоритма Random Forest (RF) // Инновации. Наука. Образование. — 2021. — №35. — С.33-39. [Nosova GS, Abdullin AH. Machine learning based on nonparametric and nonlinear Random Forest (RF) algorithm. Innovation. The science. Education. 2021; 35: 33-39. (In Russ.)]
25. Scikit-learn. Random forest, regressor. Available at: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html>. Accessed 10.10.2023.
26. Zhang И, Ren J, Wei Z, et al. Health data driven on continuous blood pressure prediction based on gradient boosting decision tree algorithm. IEEE Access. 2019; 7: 32423-32433. doi: 10.1109/ACCESS.2019.2902217.
27. Plaia A, Buscemi S, Fürnkranz J, Mencía EL. Comparing Boosting and Bagging for Decision Trees of Rankings. Journal of Classification. 2022; 39(1): 78-99. doi: 10.1007/s00357-021-09397-2.

АСТАНИН П.А.,

ФГАОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России, Москва, Россия, e-mail: med_cyber@mail.ru

РОНЖИН Л.В.,

ФГАОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России, Москва, Россия, e-mail: levronzhin@gmail.com

РАУЗИНА С.Е.,

к.м.н., доцент, ФГАОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России, Москва, Россия,
e-mail: rauzina@mail.ru

АЛГОРИТМ ОЦЕНКИ СПЕЦИФИЧНОСТИ ТЕРМИНОВ МЕТАТЕЗАУРУСА UMLS НА ПРИМЕРЕ АНАЛИЗА СЕМАНТИЧЕСКОЙ МОДЕЛИ ДЛЯ ДИФФЕРЕНЦИАЛЬНОЙ ДИАГНОСТИКИ АКСИАЛЬНОГО СПОНДИЛОАРТРИТА

DOI: 10.25881/18110193_2023_3_30

Аннотация. Актуальность. Ранняя диагностика аксиального спондилоартрита (аксСпА) является актуальной задачей, требующей поддержки принятия клинических решений (ППКР). В настоящее время значительная часть данных, применимых для разработки систем ППКР, представлена в неструктурированном виде. Семантический анализ медицинских текстов – сложная задача, требующая создания универсальных инструментов для извлечения именованных сущностей. Крупнейшим источником данных для аннотирования биомедицинских текстов является Unified Medical Language System (UMLS) – международный метатезаурус, включающий свыше 11 млн вариантов написания 4,6 млн терминов (концептов). Ключевой проблемой при использовании UMLS для анализа медицинских текстов является наличие большого количества неспецифичных (общих) терминов, не имеющих явного клинического смысла. Применение таких концептов приводит к значительному ухудшению результатов поиска, что указывает на необходимость создания инструментов автоматической оценки степени специфичности терминов UMLS.

Цель исследования. Разработка алгоритма для оценки степени специфичности терминов метатезауруса UMLS (на примере аксиального спондилоартрита).

Методы и материалы. В качестве источника информации для автоматического поиска клинических терминов использовались аннотации к англоязычным научным статьям. Тексты аннотаций извлекались с применением средств поисковой системы PubMed и помещались в единый электронный корпус, использованный для последующего извлечения концептов метатезауруса UMLS. Каждый из 24276 текстов корпуса имел однозначную метку принадлежности к одному из заболеваний дифференциального ряда для аксСпА. В общий свод включено 8260 концептов, каждый из которых получил экспертную бинарную метку относительной специфичности.

Результаты. Сформирован набор правил, основанных на сравнении средних длин иерархических цепей атомарных формулировок терминов, общего числа прямых связей, TF-IDF меры и числа связей «родитель-потомок» концептов UMLS. Данные правила включены в итоговый алгоритм оценки специфичности концептов, точность которого при попарном сравнении составила 99,1% для тестовой выборки. Однако точность модели при бинарной классификации всех концептов из выделенного свода терминов составила 74,2%, что является недостаточным для обоснования его применения при автоматическом сокращении терминологических сводов большого объема. Сформированы критерии и ограничения для использования разработанного алгоритма в процессе аннотации клинических документов.

Ключевые слова: UMLS, метатезаурус, семантический анализ текста, граф, анализ естественного языка, аксиальный спондилоартрит.

Для цитирования: Астанин П.А., Ронжин Л.В., Раузина С.Е. Алгоритм оценки специфичности терминов метатезауруса UMLS на примере анализа семантической модели для дифференциальной диагностики аксиального спондилоартрита. *Врач и информационные технологии.* 2023; 3: 30-43. doi: 10.25881/18110193_2023_3_30.

ASTANIN P.A.,

Pirogov Russian National Research Medical University, Moscow, Russia, e-mail: med_cyber@mail.ru

RONZHIN L.V.,

Pirogov Russian National Research Medical University, Moscow, Russia, e-mail: levronzhin@gmail.com

RAUZINA S.E.,

PhD, Associate Professor, Pirogov Russian National Research Medical University, Moscow, Russia, e-mail: rauzina@mail.ru

ALGORITHM FOR UMLS METATHESAURUS CONCEPTS SPECIFICITY ESTIMATION USING EXAMPLE OF ANALYSIS OF THE SEMANTIC MODEL DESCRIBING AXIAL SPONDYLOARTHRITIS DIFFERENTIAL DIAGNOSTICS

DOI: 10.25881/18110193_2023_3_30

Abstract. *Background.* Early axial spondyloarthritis (axSpA) diagnostics is a difficult task requiring clinical decision support (CDS) making. Currently, there is a big unstructured data applicable in CDS systems development. Semantic data analysis is a complex issue to solve, and unified tools for named entity recognition are required. The biggest data source for biomedical text annotation is the Unified Medical Language System (UMLS) Metathesaurus. It includes more than 11 million atomic terms for writing of 4.6 million concepts. The main issue in UMLS using for medical text analysis is a presence of numerous unspecified (generic) terms without any clinical value. Their application leads to significant decrease of searching results. That is why tools for automatic specificity degree estimation are needed to be developed.

Aim. To develop an algorithm for specificity degree estimation for UMLS metathesaurus concepts (using example of axial spondyloarthritis).

Methods. English clinical abstracts have been used as data source for automatic UMLS named entity recognition. They have been extracted using free search engine PubMed followed by integration into single electronic corpus. Then each of 24276 texts in corpus has been labeled (affiliated with one of diagnosis in differential list for axSpA) and used for UMLS concepts mapping. A total of 8260 UMLS concepts have been recognized. Each term received an expert binary label of relative specificity.

Results. Rules for concepts specificity degree estimation have been developed based on comparison of 4 parameters: mean length of hierarchical chain, total count of direct relationships, TF-IDF score and count of hierarchical relationships with child concepts UMLS. These rules have been integrated into the total algorithm for UMLS concepts specificity degree estimation. Its accuracy was 99,1% for test data sample for paired comparisons. But its accuracy for solid comparison of all extracted concepts was 74,2%, which less than desirable for substantiation of this algorithm use for automatically terms big sets cutbacks. That is why some limitations for developed algorithm have been outlined.

Keywords: UMLS, metathesaurus, semantics, graph, natural language processing, axial spondyloarthritis.

For citation: Astanin P.A., Ronzhin L.V., Rauzina S.E. Algorithm for UMLS metathesaurus concepts specificity estimation using example of analysis of the semantic model describing axial spondyloarthritis differential diagnostics. Medical doctor and information technology. 2023; 3: 30-43. doi: 10.25881/18110193_2023_3_30.

ВВЕДЕНИЕ

Аксиальный спондилоартрит (аксСпА) — это хроническое воспалительное заболевание с преимущественным поражением позвоночника и крестцово-подвздошных суставов [1]. Важной особенностью аксСпА является вариабельность клинических признаков, приводящая к формированию обширного дифференциально-диагностического ряда [2]. Основным и наиболее частым проявлением аксСпА является боль в спине, сопровождающаяся ощущениями утренней скованности движений [3–4]. Нередко у пациентов с данным заболеванием могут наблюдаться внеаксиальные и внескелетные проявления [5–6]. К внеаксиальным поражениям относятся патогенетически связанные с аксСпА заболевания опорно-двигательного аппарата: периферические артриты и энтезиты [7]. Наиболее распространенными внескелетными проявлениями являются увеиты, поражения кожи, а также воспалительные заболевания кишечника, сердца и сосудов [8–10].

Выбор тактики ведения пациентов с неустановленным диагнозом аксСпА на уровне первичного звена системы здравоохранения является сложной междисциплинарной проблемой, требующей персонифицированного подхода [11–13]. В значительной мере, своевременному выявлению аксСпА препятствуют системный характер и отсутствие патогномоничных признаков данного заболевания [14]. Период времени от начала возникновения первых клинических проявлений аксСпА до окончательного подтверждения диагноза, по разным оценкам, составляет 5–8 лет [15]. Определенные надежды на повышение качества ранней диагностики аксСпА возлагают на разработку и внедрение систем поддержки принятия клинических решений (СППКР).

Создание СППКР для дифференциальной диагностики системных заболеваний требует использования больших данных, в полной мере описывающих широкое признаковое пространство [16–18]. Значительная часть клинических данных, применимых при разработке подобных систем, хранится в электронных медицинских картах и представлена в неструктурированном (текстовом) виде [19–20]. Ключевой этап разработки СППКР, основанных на анализе неструктурированного текста, заключается в

создании терминологического свода, описывающего изучаемую клиническую область [21–24]. Обязательным требованием служит строгое соответствие используемых понятий общепринятым согласованным номенклатурам [25–26].

В настоящее время крупнейшим систематизированным сводом терминов, предназначенным для описания различных областей биологии и медицины, является Unified Medical Language System (UMLS) [27]. Актуальная версия метатезауруса UMLS (2022AB) содержит свыше 4,6 млн концептов — клинических и параклинических понятий, представленных в 76 справочниках. Каждый концепт UMLS имеет собственные варианты написания, а также внутреннюю систему приоритетов и статусов терминов. Общее число англоязычных и русскоязычных атомарных формулировок написания для всех концептов UMLS превышает 11,2 млн.

Практически каждый концепт UMLS связан как минимум с одним другим концептом, что позволяет представить данный свод терминов в виде гигантского ориентированного мультиграфа [28]. Разработка инструментов автоматического анализа графовой информационной модели UMLS внесет значительный вклад в создание неспецифичных универсальных средств обработки естественного языка (NLP). В свою очередь, использование технологий NLP обеспечит возможность решения клинических задач с использованием анализа неструктурированных текстов — основной формы представления и хранения данных в современных медицинских информационных системах [29–31].

Анализ метатезауруса UMLS — сложная и нетривиальная задача. Основной проблемой, связанной с извлечением терминов UMLS из текста, является наличие обобщающих понятий, вносящих существенный вклад в ухудшение результатов поиска. Разработка алгоритма оценки специфичности концептов метатезауруса UMLS позволит создать инструменты автоматической элиминации обобщающих терминов при извлечении именованных сущностей из текста. Использование указанных инструментов позволит значительно улучшить качество извлечения релевантной информации из графовой модели UMLS.

Целью настоящего исследования является разработка алгоритма оценки относительной специфичности концептов метатезауруса UMLS

(на примере терминов для дифференциальной диагностики аксиального спондилоартрита).

МАТЕРИАЛ И МЕТОДЫ

Исследование проведено в рамках программы стратегического академического лидерства «Приоритет — 2030» на базе Института цифровой трансформации медицины (ИЦТМ) ФГАОУ ВО «Российский национальный исследовательский медицинский университет имени Н.И. Пирогова» Минздрава России. Используемый дифференциально-диагностический ряд для аксСпА включал 9 заболеваний костно-мышечной системы: стеноз поясничного отдела позвоночного канала (M48.0, M99.5–M99.7), спондилолистез поясничного отдела (M43.1), спинальная нестабильность (M53.2), миофасциальный болевой синдром (M79.1), инфекционные поражения структур позвоночника (M49.0–M49.2, M49.3, M86), новообразования в области поясничного и крестцового отделов позвоночника (D16.6, D16.8), болезнь Форестье (M48.1), болезнь Педжета (M88) и анкилозирующий спондилит (M45) [32–35].

Для извлечения именованных сущностей применялись международные клинические справочники, представленные в актуальной версии метатезауруса UMLS (2022AB). В качестве источника информации для автоматического поиска концептов использованы тексты 24276 аннотаций к англоязычным статьям по исследуемой клинической области. При формировании корпуса текстов для всех кодов МКБ-10, заявленных для перечисленных ранее заболеваний, извлекались все англоязычные

формулировки из UMLS. Затем в поисковой системе PubMed автоматически создавались запросы с использованием в качестве ключевых слов полученных вариантов написания вышеперечисленных нозологий. На основании привязки к ключевым словам в PubMed каждый извлекаемый текст аннотации получал собственную метку связи с определенным заболеванием из дифференциального ряда для аксСпА. Тексты аннотаций сохранялись на жесткий диск, подвергались предобработке и извлечению именованных сущностей. Подробное описание данной части исследования дано в одной из предыдущих работ [36].

В ходе обработки всего корпуса текстов сформированы пересекающиеся своды терминов, описывающих клинические особенности заболеваний из дифференциального ряда. В общей сложности в созданную номенклатуру понятий включено 8260 концептов UMLS. Доля понятий, присущих только одному заболеванию (условно-патогномоничных в рамках дифференциального-диагностического ряда), составила ~41% от общего числа извлеченных терминов.

Каждому извлеченному термину экспертным способом присвоена бинарная метка относительной специфичности. В зависимости от значения данной метки были сформированы две статистические группы терминов. В первую группу включены относительно специфичные (частные) термины, во вторую — относительно неспецифичные (общие). Примеры пар относительно частных и общих терминов представлены в таблице 1.

Таблица 1 — Примеры пар относительно неспецифичных и относительно специфичных терминов UMLS

№	Неспецифичный (общий) термин	Специфичный (частный) термин
1	Боль в спине неуточненной локализации	Боль в пояснице
2	Артрит	Артрит крестцово-подвздошного сустава
3	Симптом	Симптомы заболеваний желудочно-кишечного тракта
4	Системные заболевания	Болезнь Рейтера
5	Заболевания мягких тканей	Инфекции мягких тканей
6	Нарушение баланса электролитов	Гипокалиемия
7	Некроз	Сухая гангрена стопы
8	Неуточненное заболевание молочной железы	Мастопатия
9	Отсутствие признака	Отсутствие аппетита
10	Признаки	Признаки поражения кожи
11	HLA-антигены	Положительный результат теста на HLA-B27
12	Аксиальный спондилоартрит	Анкилозирующий спондилит

Сформулирован перечень из пяти количественных параметров для сравнительной оценки степени специфичности терминов UMLS. Первый параметр отражает плотность структуры связей концепта в графовой модели и определяется как количество прямых связей между исследуемым и соседними узлами.

Второй параметр характеризует положение концепта в иерархии графовой модели UMLS. Расчет значений данного параметра осуществляется с использованием специальной таблицы Mrhier оригинального метатезауруса. Указанная таблица содержит информацию об источнике концепта, его положении в локальной иерархии терминов, формулировке родительского термина и уточнении связей. Каждая запись в таблице Mrhier позволяет определить, к какой иерархической цепочке привязана отдельно взятая формулировка UMLS. Положение концепта в иерархии графовой модели метатезауруса рассчитывается как среднее значение длин иерархических цепей (или числа связей), соединяющих корневое понятие с конечным концептом, включающим атомарные формулировки соответствующего понятия. Простейший пример иерархических цепей для термина

«боль в нижней части спины» представлен на рисунке 1.

Согласно данным, представленным на рисунке 1, термин «боль в нижней части спины» является конечным элементом трех иерархических цепей. Так, для цепи с корневым термином «симптомы и признаки» количество связей в иерархической цепи равно трём. Для цепи с корневым термином «Анатомические образования» длина иерархической цепи будет равна двум. Для цепи с корневым термином «Клинические находки» длина цепи будет равна единице. Таким образом, среднее значение длин иерархических цепей для термина «боль в нижней части спины» равняется двум. Подобным образом значения количественной меры, отражающей положение концепта в иерархической структуре, можно рассчитать для всех терминов UMLS.

Третьим параметром для оценки специфичности терминов UMLS является рассчитанное значение TF-IDF меры. Математический смысл TF-IDF меры состоит в том, что вес токена (отдельного слова, фразы или термина) прямо пропорционален TF (term frequency) — частоте употребления в документе и обратно пропорционален IDF (inverse document frequency) — частоте

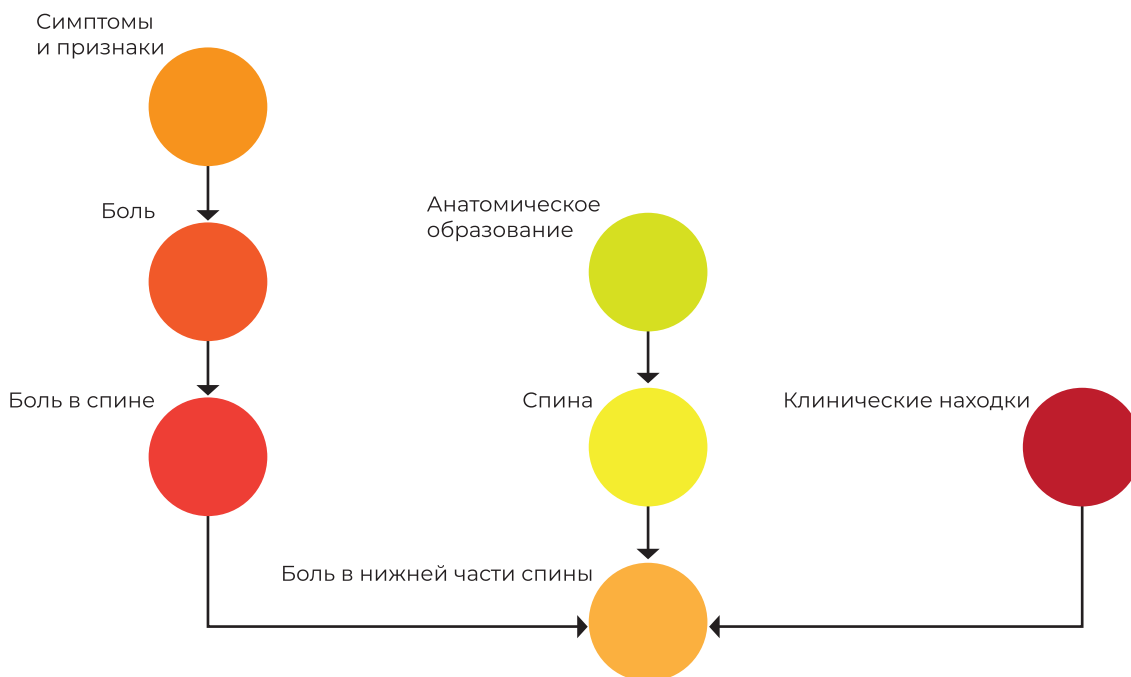


Рисунок 1 — Примеры иерархических цепей для термина «боль в нижней части спины».

использования во всех документах корпуса [37]. TF-IDF мера, использованная в данном исследовании для определения степени специфичности терминов UMLS, рассчитывалась по следующей формуле:

$$TF - IDF_i = \frac{n_i}{m_i} \cdot \ln \frac{D}{d_i} \quad (1),$$

где: n_i — количество токенов i в корпусе документов; m_i — общее число токенов в документах, в которых токен i встретился хотя бы один раз; D — общее количество документов в корпусе; d_i — количество документов, в которых токен i встретился хотя бы один раз.

Основным источником информации, необходимой для расчета значений TF-IDF, является NLM (U. S. National Library of Medicine) — Национальная Медицинская Библиотека США. Одним из важнейших инструментов, разработанных NLM, является SemRep (Semantic Repository) — средство для автоматического извлечения связей между концептами UMLS из неструктурированного текста. Одним из результатов работы SemRep является свободно распространяемая база данных SemMedDB (Semantic Medline Database), созданная на основе машинной обработки массивных корпусов текстов на английском языке. Данные, использованные в настоящем исследовании для определения значений TF-IDF мер, содержались в базе данных SemMedDB в таблице Predication. Каждой записи в указанной таблице сопоставлен уникальный номер публикации (PMID), служащей в качестве источника извлеченной связи. Для расчета TF-IDF мер агрегированы записи со значениями идентификаторов корневого и концевых концептов, а также PMID соответствующих англоязычных статей из поисковой системы PubMed.

Четвертый параметр позволяет оценить количество дочерних терминов изучаемого концепта. Для расчета данного параметра определяется число прямых связей, относящихся к типам CHD — child relationships (связи родительских терминов с дочерними) и RN — narrower relationships (связи более широких терминов с более узкими) [36]. Предполагается, что указанные типы вертикальных иерархических связей соединяют более общие и неспецифичные термины с более частными и специфичными.

Пятый параметр определяет семантическую сложность термина. В качестве значения данного параметра в настоящем исследовании использовалось среднее число слов, содержащихся во всех вариантах написания концепта. Параметр определялся, исходя из предположения, что более специфичные термины содержат больше языковых единиц.

Первичный анализ качества выделенных параметров осуществлялся с использованием методов непараметрической статистики. Для оценки типа распределения количественных признаков применялся критерий Колмогорова-Смирнова. Однако, поскольку вид распределения всех проверяемых признаков отличался от нормального, для описания количественных данных рассчитывались медиана и межквартильный размах (Me [Q_1 ; Q_3]). Для оценки структуры исследуемых групп определялись доли и ошибки долей с последующим выражением в процентах ($p \pm M\%$).

Для валидации алгоритма оценки специфичности терминов UMLS и его составных частей использованы дополнительные данные, представленные 424 предварительно размеченными парами концептов, не пересекающихся с номенклатурой терминов для описания заболеваний из дифференциально-диагностического ряда для аксСпА. Основным условием для формирования пар было наличие пересечений хотя бы по одному из токенов. В качестве примера пары терминов, подходящих для ранжирования по степени специфичности, можно выделить «неуточненную боль в области живота» и «боль в правом боку» с общим токеном «боль». В каждой паре тремя независимыми экспертами выделено по одному относительно специфичному и относительно неспецифичному концепту. Было предложено включать пары терминов, не подлежащих ранжированию, однако ни один из экспертов не посчитал необходимым элиминировать хотя бы одну из пар в автоматически сформированном своде.

Качество экспертных оценок подтверждалось с применением непараметрического критерия χ^2 Пирсона. Количественная оценка согласованности экспертных мнений производилась с использованием коэффициента ассоциации по следующей формуле:

$$K_a = \frac{a \cdot d - b \cdot c}{a \cdot d + b \cdot c} \quad (2),$$

где: a — число истинно положительных оценок, b — число ложноположительных оценок, c — число ложноотрицательных оценок, d — число истинно отрицательных оценок.

Экспертные оценки, полученные в ходе семантического анализа пар концептов, характеризовались высокой степенью согласованности (таблица 2). Итоговое решение об используемом варианте разметки пар терминов UMLS принималось по принципу наибольшего количества голосов.

Анализ прогностической способности правил, основанных на использовании выделенных параметров для сравнительной оценки степени специфичности терминов, осуществлялся путем вычисления стандартных метрик бинарной классификации: точности, чувствительности и специфичности. Статистически значимыми считались результаты проверки гипотез при уровне значимости $p < 0,05$.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

При оценке межгрупповых различий по параметрам для определения степени

специфичности терминов UMLS в экспертно размеченных группах статистически значимые различия выявлены лишь по четырём из пяти в сформированном перечне (таблица 3).

Согласно данным, представленным в таблице 3, медиана общего числа прямых связей концептов UMLS оказалась выше в группе относительно неспецифичных терминов. Схожая тенденция наблюдалась по числу иерархических связей, соединяющих родительские термины с дочерними. Медиана по средним значениям длин иерархических цепей оказалась выше в группе относительно специфичных терминов. Аналогичная закономерность характерна для TF-IDF меры. Все полученные закономерности оправдали изначальные ожидания и подтвердили возможность использования данных параметров для оценки специфичности терминов UMLS. Крайне важно отметить, что не были выявлены статистически значимые различия по среднему числу слов в атомарных формулировках концептов. Вопреки ожиданиям, данный параметр не продемонстрировал наличия потенциальных статистических закономерностей и,

Таблица 2 — Матрица согласованности экспертных мнений (K_a)

Номер эксперта	1	2	3
1		$K_a = 0,730; p < 0,001$	$K_a = 0,744; p < 0,001$
2	$K_a = 0,730; p < 0,001$		$K_a = 0,803; p < 0,001$
3	$K_a = 0,744; p < 0,001$	$K_a = 0,803; p < 0,001$	

Таблица 3 — Оценка межгрупповых различий по параметрам для определения степени специфичности терминов UMLS

Краткое описание параметра	Группа терминов UMLS (Me [Q1; Q3])		P
	Относительно неспецифичные (общие) термины (n = 4099, 49,6%)	Относительно специфичные (частные) термины (n = 4161, 50,4%)	
Общее количество прямых связей концепта	52,0 [35,0; 158]	44,0 [24,0; 121]	0,026
Среднее значение длин иерархических цепей атомарных формулировок концепта	4,60 [4,00; 5,83]	5,95 [4,74; 7,37]	<0,001
Значение TF-IDF меры для концепта	1,80 [1,41; 2,33]	1,97 [1,57; 2,43]	<0,001
Количество иерархических связей «родительский термин – дочерний термин» концепта	15,0 [2,00; 81,0]	12,0 [3,00; 42,0]	<0,001
Среднее число слов в атомарных формулировках концепта	2,43 [1,85; 2,97]	2,42 [1,89; 3,04]	0,468

связи с этим, не рассматривался при последующем анализе.

На основании результатов из таблицы 3 сформулированы бинарные правила для оценки степени относительной специфичности терминов UMLS. При анализе качества выделенных правил определялись следующие количественные характеристики: точность, чувствительность, специфичность, а также доля концептов, для которых является возможным применение соответствующего правила. Метрики бинарной классификации рассчитывались для набора размеченных экспертным способом пар терминов. Доли концептов, для которых возможно применение правил, рассчитывались для всего метатезауруса UMLS.

Итоговый алгоритм оценки специфичности терминов UMLS представлен на рисунке 2.

Результат проверки качества отдельных правил и итогового алгоритма представлены в таблице 4.

Из данных таблицы 4 следует, что наилучшими прогностическими характеристиками для выявления относительно специфичных и относительно неспецифичных концептов обладают правила, основанные на сравнении общего числа прямых связей и числа связей «родитель-потомок». Важно отметить, что наиболее применимыми для UMLS являются правила, подразумевающие оценку общего числа прямых связей и среднюю длину иерархических цепей формулировок сравниваемых концептов. Точность итогового алгоритма составила 99,1% для тестовой выборки.

Предпринята попытка бинарной классификации терминов в исходно сформированном

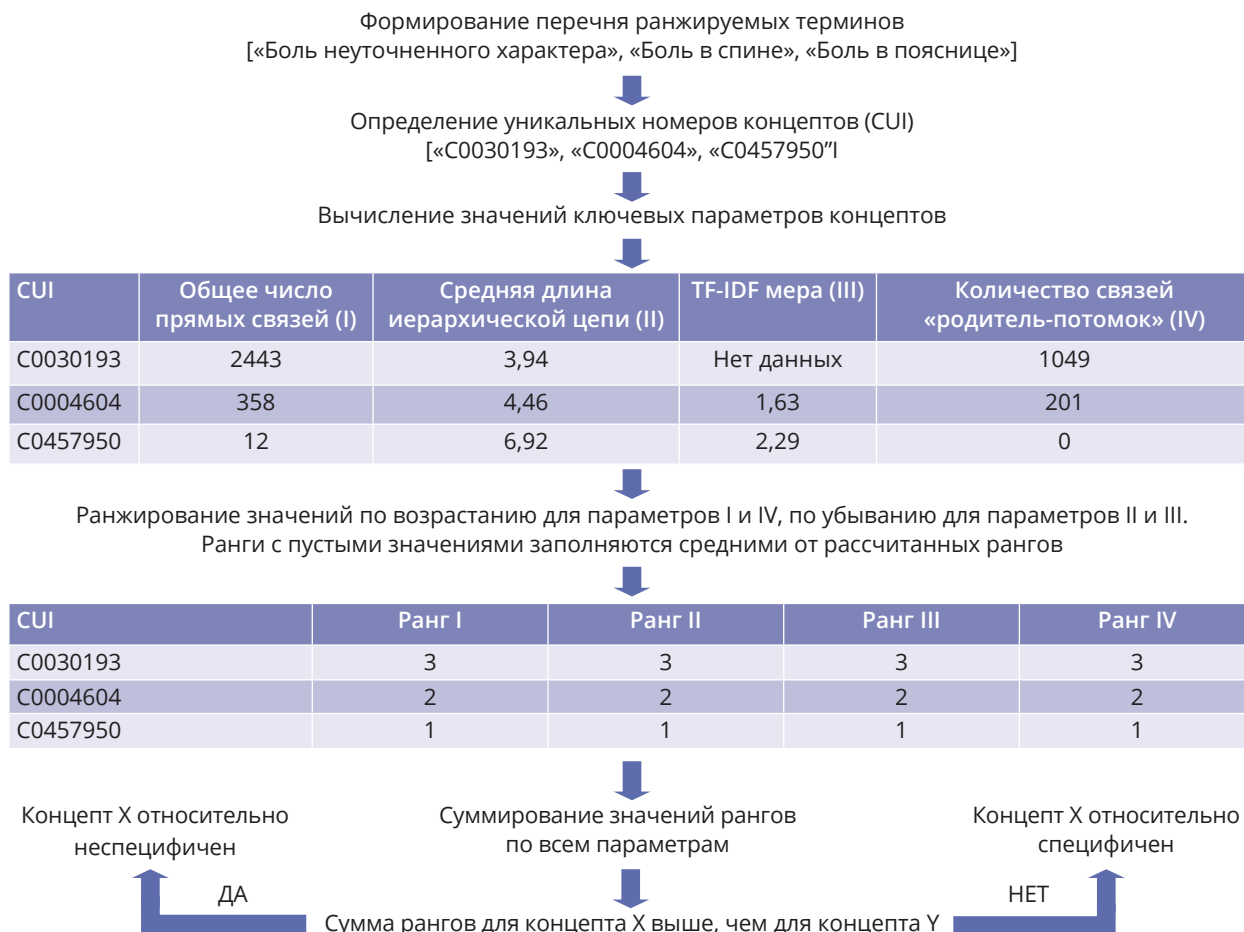


Рисунок 2 — Алгоритм оценки относительной специфичности терминов метатезауруса UMLS.

Таблица 4 — Правила для оценки степени относительной специфичности терминов метатезауруса UMLS

№	Логическое правило	Метрики качества бинарной классификации, %			Число и доля (n, %) концептов, для которых возможно применение правила
		А	Б	В	
1	Количество прямых связей неспецифичного концепта больше, чем у специфичного	97,2	97,7	96,6	4548967 (98,6%)
2	Длина иерархической цепи неспецифичного концепта меньше, чем у специфичного	91,3	92,7	89,8	3600628 (78,1%)
3	TF-IDF мера неспецифичного концепта меньше, чем у специфичного	83,5	85,8	81,1	300459 (6,51%)
4	Количество связей «родитель-потомок» неспецифичного концепта больше, чем у неспецифичного	97,9	97,7	98,1	864833 (18,8%)
5	Итоговый алгоритм	99,1	99,1	99,0	864833 (18,8%)

Примечание: А — Точность, Б — Чувствительность, В — Специфичность.

терминологическом своде для описания клинических особенностей заболеваний из дифференциально-диагностического ряда для аксСпА. Точность модели при бинарной классификации всех 8260 концептов составила 74,2%, что является недостаточным для обоснования его применения при автоматическом сокращении терминологических сводов большого объема.

Таким образом, внедрение разработанного алгоритма для использования при аннотации клинических документов требует предварительного формирования перечня критериев и ограничений. Одним из таких критериев может стать наличие пересечений токенов в извлеченных терминах. Также допускается использование алгоритма для ранжирования небольших наборов терминов UMLS, связанных прямыми ассоциативными связями.

ОБСУЖДЕНИЕ

Автоматическое аннотирование неструктурированных медицинских текстов представляет большой практический интерес для современной системы здравоохранения [38]. На сегодняшний день крупнейшим источником биомедицинских терминов, пригодных для использования при обработке клинических текстов, является метатезаурус UMLS. Однако процесс извлечения именованных сущностей с применением UMLS сопровождается ключевой проблемой, связанной с агрегацией избыточных и

обобщающих понятий, не относящихся к признакам исследуемых заболеваний. В нашей работе доля таких терминов составила около 50%. Использование подобных терминов при решении информационно-поисковых задач без предварительной обработки и ранжирования приводит к значительному ухудшению качества извлекаемой информации.

В настоящем исследовании была предпринята попытка создания универсального алгоритма для ранжирования терминов по степени относительной специфичности. Выделены основные параметры, используемые при создании правил ранжирования и базируемые на анализе семантических характеристик терминов UMLS. Необходимо отметить, что наименьшую эффективность продемонстрировал параметр, основанный на оценке среднего числа слов, входящих в термины исследуемого концепта. Данный факт указывает на низкое качество отбора специфичных концептов при использовании правил, основанных только на подстрочном поиске. Остальные параметры, описанные в настоящей работе (общее количество прямых связей, средняя длина иерархической цепи, TF-IDF мера и количество связей «родитель-потомок») демонстрировали высокую точность и были использованы при создании итогового алгоритма. При попарном сравнении концептов точность данного алгоритма составила 99,1%, однако по мере увеличения числа ранжируемых концептов

качество оценки относительной специфичности заметно снижается из-за увеличения накопленной ошибки. В связи с этим применение разработанного алгоритма для элиминации обобщающих терминов из сводов большого объема не представляется возможным. Предполагается, что максимальная эффективная работа алгоритма будет достигнута при ранжировании небольших наборов терминов. В настоящее время разработанный алгоритм внедрен в аналитические сервисы информационно-поисковой системы, разрабатываемой на базе Института цифровой трансформации медицины РНИМУ им. Н.И. Пирогова.

ЗАКЛЮЧЕНИЕ

Разработанный алгоритм для оценки относительной специфичности терминов метатезауруса UMLS продемонстрировал высокую точность (99,1%) при попарном сравнении концептов с пересечением токенов, что указывает на широкие возможности использования данного

алгоритма при аннотировании клинических текстов. Одним из перспективных направлений для продолжения настоящего исследования является создание системы поддержки принятия решений для дифференциальной диагностики аксиального спондилоартрита с использованием анализа неструктурированного текста.

Конфликт интересов: Авторы заявляют об отсутствии конфликта интересов.

Финансирование: Исследование выполнено в рамках федеральной программы «Приоритет 2030».

Благодарность. Авторы выражают признательность сотрудникам Института цифровой трансформации медицины (ИЦТМ), Ядгаровой П.А., Николаиди Е.Н., Потаповой И.И., за экспертную работу по разметке валидационного набора данных. Авторы выражают отдельную благодарность директору ИЦТМ, Зарубиной Т.В., за помощь в общей корректуре и редактуре текста данной статьи.

ЛИТЕРАТУРА/REFERENCES

1. Моисеев С.В., Новиков П.И., Гуляев С.В. и др. Анкилозирующий спондилит: подходы к диагностике и клиническая эффективность упадацитиниба // Клиническая фармакология и терапия. — 2021. — Т.30. — №4. — С.62-70. [Moiseev S, Novikov P, Gulyaev S, et al. Ankylosing spondylitis: diagnostic challenges and efficacy of upadacitinib. *Klinicheskaja farmakologija i terapija*. 2021; 30(4): 62-70. (In Russ.)] doi: 10.32756/0869-5490-2021-4-62-70.
2. Кричевская О.А., Дубинина Т.В., Ильиных Е.В. и др. Боль в спине и оценка активности анкилозирующего спондилита на фоне гестации: симптомы, отражающие обострение заболевания, и беременность // Современная ревматология. — 2022. — Т.16. — №5. — С.60-65. [Krichevskaya OA, Dubinina TV, Ilinykh EV, et al. Back pain and assessment of ankylosing spondylitis activity during gestation: symptoms reflecting exacerbation of the disease and pregnancy. *Modern Rheumatology Journal*. 2022; 16(5): 60-65. (In Russ.)] doi: 10.14412/1996-7012-2022-5-60-65.
3. Дубинина Т.В., Демина А.Б. Методы лучевой диагностики как инструмент мониторинга аксиального спондилоартрита в реальной клинической практике // Современная ревматология. — 2022. — Т.16. — №1. — С.91-96. [Dubinina TV, Demina AB. Radiologic methods as a tool for monitoring axial spondyloarthritis in real clinical practice. *Modern Rheumatology Journal*. 2022; 16(1): 91-96. (In Russ.)] doi: 10.14412/1996-7012-2022-1-91-96.
4. Варавин Н.А., Верткин А.Л. Боль в спине в терапевтической практике // Лечащий врач. — 2022. — Т.25. — №7-8. — С.52-56. [Varavin NA, Vertkin AL. Back pain in therapeutic practice. *The Attending Physician*. 2022; 25(7-8): 52-56. (In Russ.)] doi: 10.51793/OS.2022.25.8.008.
5. Каратеев Д.Е., Степанова Е.А., Лучихина Е.Л. Методические рекомендации по рентгенологическим методам исследования при ревматоидном артрите и анкилозирующем спондилите // Эффективная фармакотерапия. — 2022. — Т.18. — №18. — С.12-25. [Karateev DE, Stepanova EA, Luchikhina EL. Practical guidelines for radiological investigation methods in rheumatoid arthritis and ankylosing spondylitis. *Effective Pharmacotherapy*. 2022; 18(18): 12-25. (In Russ.)] doi: 10.33978/2307-3586-2022-18-18-12-25.

6. Гайдук А.С., Железняк И.С., Тыренко В.В. и др. Цифровой томосинтез и другие методы визуализации в ранней диагностике аксиальных спондилоартритов: обзор литературы // Лучевая диагностика и терапия. — 2022. — №2(13). — С.25-35. [Gaiduk AS, Zheleznyak IS, Tyrenko VV, et al. Digital tomosynthesis and other visualization methods in the early detection of axial spondyloarthritis: a review. Diagnostic radiology and radiotherapy. 2022; 2(13): 25-35. (In Russ.)] doi: 10.22328/2079-5343-2022-13-2-25-35.
7. Гараева А.Р., Лапшина С.А., Анисимов В.И. и др. Клинико-инструментальная диагностика ранних изменений атлантаксиальной области при анкилозирующем спондилите // Практическая медицина. — 2023. — Т.21. — №3. — С.68-73. [Garaeva AR, Lapshina SA, Anisimov VI, et al. Clinical and instrumental diagnostics of early changes in the atlantoaxial area in ankylosing spondylitis. Practical medicine. 2023; 21(3): 68-73. (In Russ.)] doi: 10.32000/2072-1757-2023-3-68-73.
8. Иванова Л.В., Акулинушкина Е.Ю., Лапшина С.А., Абдулганиева Д.И. Ранняя диагностика воспалительных заболеваний кишечника у пациентов со спондилоартритами // Практическая медицина. — 2023. — Т.21. — №2. — С.54-57. [Ivanova LV, Akulinushkina EYU, Lapshina SA, Abdulganieva DI. Early diagnosis of inflammatory bowel diseases in patients with spondyloarthritis. Practical medicine. 2023; 21(2): 54-57. (In Russ.)] doi: 10.32000/2072-1757-2023-2-54-57.
9. Моисеев С.В., Буланов Н.М. Аутоиммунитет, аутовоспаление и почки // Клиническая фармакология и терапия. — 2022. — Т.31. — №4. — С.7-17. [Moiseev S, Bulanov N. Autoimmunity, autoinflammation and kidney. Klinicheskaya farmakologiya i terapiya. 2022; 31(4): 7-17. (In Russ.)] doi: 10.32756/0869-5490-2022-4-7-17.
10. Пономарева М.Н., Карпова Д.А., Петров И.М. Анкилозирующий спондилит: гипотезы патогенеза, новые биомаркеры и особенности терапии // Современные проблемы науки и образования. — 2021. — №6. — С.188. [Ponomareva MN, Karpova DA, Petrov IM. Ankylosing spondylitis: hypotheses of pathogenesis, new biomarkers and features of therapy. Modern problems of science and education. 2021; 6: 188. (In Russ.)] doi: 10.17513/spno.31264.
11. Мартюшев-Поклад А.В., Гулиев Я.И., Казаков И.Ф. и др. Персонализированные инструменты цифровой трансформации здравоохранения: пути совершенствования // Врач и информационные технологии. — 2021. — №55. — С.4-13. [Martyushev-Poklad AV, Guliev YI, Kazakov IF, et al. Person-centered instruments in digital transformation of healthcare: ways to improve. Medical doctor and information technologies. 2021; 55: 4-13. (In Russ.)] doi: 10.25881/18110193_2021_55_4.
12. Батудаева Т.И., Павлова А.Б., Лобышева Е.А., Арзуманян Э.А. Анализ стационарной помощи пациентам с анкилозирующим спондилитом // Вестник Бурятского государственного университета. Медицина и фармация. — 2022. — №1. — С.7-14. [Batudaeva TI, Pavlova AB, Lobysheva EA, Arzumanyan EA. Analysis of hospital care for patients with ankylosing spondylitis. Vestnik buryatskogo gosudarstvennogo universiteta. Meditsina i farmatsiya. 2022; 1: 7-14. (In Russ.)] doi: 10.18101/2306-1995-2022-1-7-14.
13. Шостак Н.А., Правдюк Н.Г., Новикова А.В. Боль в спине у молодых: алгоритм ведения в практике врача первичного звена // Клиницист. — 2022. — Т.16. — №3. — С.48-57. [Shostak NA, Pravdyuk NG, Novikova AV. Back pain in young people: algorithm of management in practice of primary physician. Analysis of hospital care for patients with ankylosing spondylitis. The Clinician. 2022; 16(3): 48-57. (In Russ.)] doi: 10.17650/1818-8338-2022-16-3-K674.
14. Лила А.М., Дубинина Т.В., Древалъ Р.О. и др. Медико-социальная значимость и расчет экономического бремени аксиального спондилоартрита в Российской Федерации // Современная ревматология. — 2022. — Т.16. — №1. — С.20-25. [Lila AM, Dubinina TV, Dreval RO, et al. Medical and social significance and calculation of the economic burden of axial spondyloarthritis in the Russian Federation. Modern Rheumatology Journal. 2022; 16(1): 20-25. (In Russ.)] doi: 10.14412/1996-7012-2022-1-20-25.
15. Астанин П.А., Наркевич А.Н. Цифровые технологии в оценке течения заболеваний с выраженным болевым синдромом на примере анкилозирующего спондилита // Российский

- журнал боли. — 2021. — Т.19. — №2. — С.38-41. [Astaniin PA, Narkevich AN. Digital technology for estimation of course of diseases with acute pain syndrome on the example of ankylosing spondylitis. *Russian Journal of Pain*. 2021: 19(2): 38-41. (In Russ.)] doi: 10.17116/pain20211902138.
16. Киселев К.В., Ноева Е.А., Выборов О.Н. и др. Разработка алгоритма работы логического решателя интеллектуальной системы поддержки принятия врачебных решений для инструментальной диагностики стенокардии // *Медицинские технологии. Оценка и выбор*. — 2019. — №1(35). — С.32-42. [Kiselev KV, Noeva EA, Vyborov ON. Development of a reasoning solver algorithm for instrumental diagnostics of angina pectoris in intelligent clinical decision support system. *Medical Technologies. Assessment and Choice*. 2019: 1(35): 32-42. (In Russ.)] doi: 10.31556/2219-0678.2019.35.1.032-042.
 17. Кобринский Б.А., Благодосклонов Н.А., Демикова Н.С., и др. Компьютерная система для дифференциальной диагностики лизосомных болезней накопления на основе методов искусственного интеллекта // *Бюллетень сибирской медицины*. — 2022. — Т.21. — №2. — С.67-73. [Kobriniskii BA, Blagosklonov NA, Demikova NS, et al. An artificial intelligence computer system for differential diagnosis of lysosomal storage diseases. *Bulletin of Siberian Medicine*. 2022: 21(2): 67-73. (In Russ.)] doi: 10.20538/1682-0363-2022-2-67-73.
 18. Орлова Н.В., Суворов Г.Н., Горбунов К.С. Этика и правовое регулирование использования больших баз данных в медицине // *Медицинская этика*. — 2022. — Т.10. — №3. — С.4-9. [Orlova NV, Suvorov GN, Gorbunov KS. Ethics and legal regulation of using large databases in medicine. *Medical ethics*. 2022: 10(3): 4-9. (In Russ.)] doi: 10.24075/medet.2022.056.
 19. Шарова Д.Е., Михайлова А.А., Гусев А.В. Анализ мирового опыта в регулировании использования медицинских данных для целей создания систем искусственного интеллекта на основе машинного обучения // *Врач и информационные технологии*. — 2022. — №4. — С.28-39. [Sharova DE, Mikhailova AA, Gusev AV. An analysis of global experience in regulations on the use of medical data for artificial intelligence systems development based on machine learning. *Medical doctor and information technologies*. 2022: 4: 28-39. (In Russ.)] doi: 10.25881/18110193_2022_4_28.
 20. Гусев А.В., Зингерман Б.В., Тюфилин Д.С., Зинченко В.В. Электронные медицинские карты как источник данных реальной клинической практики // *Реальная клиническая практика: данные и доказательства*. — 2022. — Т.2. — №2. — С.8-20. [Gusev AV, Zingerman BV, Tyufilin DS, Zinchenko VV. Electronic medical records as a source of real-world clinical data. *Real-World Data & Evidence*. 2022: 2(2): 8-20. (In Russ.)] doi: 10.37489/2782-3784-myrwd-13.
 21. Гурдаева Н.А. Специальная лексика современного русского языка в свете теории функционально-семантического поля // *Вестник Таганрогского государственного педагогического института*. — 2012. — №2. — С.15-19. [Gurdaeva NA. Spetsial'naya leksika sovremennogo russkogo yazyka v svete teorii funktsional'no-semanticheskogo polya. *Vestnik Taganrogskogo gosudarstvennogo pedagogicheskogo instituta*. 2012: 2: 15-19. (In Russ.)]
 22. Зарубина Т.В. *Медицинская информатика: учебник* / Под ред. Т.В. Зарубиной, Б.А. Кобринского. — Москва: ГЭОТАР-Медиа, 2022. [Zarubina TV. *Medical informatics*. 2nd ed. Moscow: GEOTAR-Media, 2022. (In Russ.)] doi: 10.33029/9704-6273-7-TMI-2022-1-464.
 23. Осмоловский И.С., Зарубина Т.В., Шостак Н.А. и др. Разработка структуры базы знаний в области диагностики подагры // *Сибирский журнал клинической и экспериментальной медицины*. — 2022. — Т.37. — №3. — С.149-158. [Osmolovsky IS, Zarubina TV, Shostak NA. Development of knowledge base structure for gout diagnosis. *The Siberian Journal of Clinical and Experimental Medicine*. 2022: 37(3): 149-158. (In Russ.)] doi: 10.29001/2073-8552-2022-37-3-149-158.
 24. Будыкина А.В., Тихомирова Е.В., Киселев К.В. и др. Формализация знаний о желудочно-кишечном кровотечении неясного генеза для использования в интеллектуальных системах поддержки принятия врачебных решений // *Вестник новых медицинских технологий*. — 2020. — Т.27. — №4. — С.98-101. [Budykina AV, Tikhomirova EV, Kiselev KV, et al. Formalization of knowledge about gastrointestinal bleeding of unknown origin for use in intelligent clinical decision support systems. *Journal of New Medical Technologies*. 2020: 27(4): 98-101. (In Russ.)] doi: 10.24411/1609-2163-2020-16741.

25. Колесникова О.И., Лопатина Е.В., Соколова В.В. Терминологические соответствия при переводе экономических терминов с английского на русский язык // Международный научно-исследовательский журнал. — 2021. — №1-3(103). — С.153-157. [Kolesnikova OI, Lopatina EV, Sokolova VV. Terminological correspondence in the translation of economic terms from English to Russian. Mezhdunarodnyi nauchno-issledovatel'skii zhurnal. 2021: 1-3(103): 153-157. (In Russ.)] doi: 10.23670/IRJ.2021.103.1.083.
26. Зацман И.М. Формы представления нового знания, извлеченного из текстов // Информатика и ее применения. — 2021. — Т.15. — №3. — С.83-90. [Zatsman IM. Forms representing new knowledge discovered in texts. Informatics and Applications. 2021: 15(3): 83-90. (In Russ.)] doi: 10.14357/19922264210311.
27. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. Nucleic Acids Res. 2004; 32: 267-270. doi: 10.1093/nar/gkh061.
28. Астанин П.А. Применение автоматизированного анализа семантической сети UMLS для решения задачи поиска релевантных знаний о ревматических заболеваниях // Математическое моделирование систем и процессов. 2022. — С.6-12. [Astaniin PA. Primenenie avtomatizirovannogo analiza semanticheskoi seti UMLS dlya resheniya zadachi poiska relevantnykh znaniy o revmaticheskikh zabolevaniyakh. Matematicheskoe modelirovanie sistem i protsessov. 2022: 6-12. (In Russ.)] doi: 10.37490/978-5-00200-102-6-6-12.
29. Тучкова П.А. Применение методов обработки естественного языка для анализа текстовых и речевых данных в медицине // Наукосфера. — 2021. — №5-1. — С.174-179. [Tuchkova PA. Application of natural language processing methods for analysing of text and speech data in medicine. Naukosfera. 2021: 5-1: 174-179. (In Russ.)] doi: 10.5281/zenodo.4771893.
30. Сбоев А.Г., Селиванов А.А., Рыбка Р.Б. Современные методы экстракции связанных именованных сущностей на примере биомедицинских текстовых данных // Вестник Военного инновационного технополиса «Эра». — 2022. — Т.3. — №1. — С.57-67. [Sboev AG, Selivanov AA, Rybka RB. Sovremennyye metody ekstraktsii svyazannykh imenovannykh sushchnostei na primere biomeditsinskikh tekstovykh dannyykh. Vestnik voennogo innovatsionnogo tekhnopolisa «Era» (In Russ.)] doi: 10.56304/S2782375X22010193.
31. Гусев А.В., Владимирский А.В., Голубев Н.А., Зарубина Т.В. Информатизация здравоохранения Российской Федерации: история и результаты развития // Национальное здравоохранение. — 2021. — Т.2. — №3. — С.5-17. [Gusev AV, Vladzimirskii AV, Golubev NA, Zarubina TV. Informatization of healthcare in the Russian Federation: history and results of development. National Health Care (Russia). 2021: 2(3): 5-17. (In Russ.)] doi: 10.47093/2713-069X.2021.2.3.5-17.
32. Никитина Н. М., Юпатова М. И., Ребров А. П. Проблемы остеопороза у пациентов с анкилозирующим спондилитом (обзор литературы) // Медицинский алфавит. — 2023. — №9. — С.40-45. [Nikitina NM, Yupatova MI, Rebrov AP. Problems of osteoporosis in patients with ankylosing spondylitis (literature review). Medical alphabet. 2023: 9: 40-45. (In Russ.)] doi: 10.33667/2078-5631-2023-9-40-45.

33. Годзенко А.А., Черемушкина Е.В., Димитрева А.Е., Урумова М.М. Сочетание анкилозирующего спондилита и ревматоидного артрита: клинические наблюдения и обзор литературы // Современная ревматология. — 2021. — Т.15. — №4. — С.72-80. [Godzenko AA, Cheremushkina EV, Dimitreva AE, Urumova MM. Combination of ankylosing spondylitis and rheumatoid arthritis: clinical observations and literature review. Modern Rheumatology Journal. 2021: 15(4): 72-80. (In Russ.)] doi: 10.14412/1996-7012-2021-4-72-80.
34. Нурполатова С.Т., Косымбетова А.Б., Джуманазарова Г.У. Боль в спине, как одна из проблем медицины // Бюллетень науки и практики. — 2021. — Т.7. — №6. — С.200-207. [Nurpolatova S, Kosymbetova A, Dzhumanazarova G. Back pain, as one of the problems of medicine. Bulletin of science and practice. 2021: 7(6): 200-207. (In Russ.)] doi: 10.33619/2414-2948/67/23.
35. Олюнин Ю.А., Лила А.М. Хроническая боль в спине глазами ревматолога // Современная ревматология. — 2022. — Т.16. — №5. — С.94-100. [Olyunin YuA, Lila AM. Chronic back pain from rheumatologist point of view. Modern Rheumatology Journal. 2022: 16(5): 94-100. (In Russ.)] doi: 10.14412/1996-7012-2022-5-94-100.
36. Астанин П.А., Раузина С.Е., Зарубина Т.В. Автоматизированная система извлечения клинически релевантных терминов UMLS из текстов англоязычных статей на примере аксиального спондилоартрита. Социальные аспекты здоровья населения. — 2023. — Т.69. — №3. — С.1. [Astaniin PA, Rauzina SE, Zarubina TV. Automated system for recognizing clinically relevant UMLS terms in texts of the English-language articles exemplified by axial spondyloarthritis. Social Aspects of Population Health. 2023: 69(3): 1. (In Russ.)] doi: 10.21045/2071-5021-2023-69-3-14.
37. Валиев А.И., Лысенкова С.А. Применение методов машинного обучения для автоматизации процесса анализа содержания текста // Вестник кибернетики. — 2021. — №4(44). — С.12-15. [Valiev AI, Lysenkova SA. Application of machine learning methods for automation of the process of the text contents analysis. Proceedings in Cybernetics. 2021: 4(44): 12-15. (In Russ.)] doi: 10.34822/1999-7604-2021-4-12-15.
38. Зулкарнеев Р.Х., Юсупова Н.И., Сметанина О.Н. и др. Методы и модели извлечения знаний из медицинских документов // Информатика и автоматизация. — 2022. — Т.21. — №6. — С.1169-1210. [Zulkarneev R, Yusupova N, Smetanina O. Method and models of extraction of knowledge from medical documents. Informatics and Automation. 2022: 21(6): 1169-1210. (In Russ.)] doi: 10.15622/ia.21.6.4.

ШАХГЕЛЬДЯН К.И.,

д.т.н., доцент, Владивостокский государственный университет, Владивосток, Россия,
e-mail: carinashakh@gmail.com

ГЕЛЬЦЕР Б.И.,

член-корр. РАН, д.м.н., профессор, Дальневосточный федеральный университет,
г. Владивосток, Россия, e-mail: boris.geltser@vvsu.ru

КУКСИН Н.С.,

Дальневосточный федеральный университет, г. Владивосток, Россия, e-mail: nikita.kuksin@vvsu.ru

ДОМЖАЛОВ И.Г.,

Дальневосточный федеральный университет, г. Владивосток, Россия, e-mail: igor@domzhalov.ru

МНОГОУРОВНЕВАЯ КАТЕГОРИЗАЦИЯ НЕПРЕРЫВНЫХ ПЕРЕМЕННЫХ В ЗАДАЧАХ ОБЪЯСНЕНИЯ ПРОГНОЗНЫХ ОЦЕНОК МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ В КЛИНИЧЕСКОЙ МЕДИЦИНЕ

DOI: 10.25881/18110193_2023_3_44

Аннотация. Цель. Сравнительная оценка качества прогностических моделей внутригоспитальной летальности (ВГЛ) у больных инфарктом миокарда с подъемом сегмента ST (ИМnST) после чрескожного коронарного вмешательства (ЧКВ), разработанных на основе предикторов в непрерывной, дихотомической и многоуровневой категориальной формах.

Материалы и методы. Проведено одноцентровое ретроспективное исследование, в рамках которого анализировали данные 4677 историй болезни пациентов с ИМnST после ЧКВ, находившихся на лечении в Региональном сосудистом центре г. Владивостока. Было выделено 2 группы больных: первая- 318 (6,8%) человек, умерших в стационаре, вторая — 4359 (93,2%) — с благоприятным исходом лечения. Прогностические модели ВГЛ с непрерывными переменными были разработаны методами многофакторной логистической регрессии, случайного леса и стохастического градиентного бустинга. Дихотомизация предикторов выполнялась методами поиска на сетке оптимальных точек отсечения, расчета центроидов и аддитивного объяснения Шепли (SHAP). Для многоуровневой категоризации предложено использовать объединение пороговых значений, выделенных при дихотомизации, а также ранжирование порогов отсечения с помощью весовых коэффициентов многофакторной логистической регрессии.

Результаты. По результатам многоступенчатого анализа показателей клинко-функционального статуса больных ИМnST были выделены и валидированы новые предикторы ВГЛ, выполнена их категоризация и разработаны прогностические модели с непрерывными, дихотомическими и многоуровневыми категориальными переменными (AUC: 0.885-0.902). Модели, предикторы которых были выделены методом мультиметрической категоризации, не уступали по точности моделям с непрерывными переменными и имели более высокие метрики качества, чем алгоритмы с дихотомическими предикторами. Преимущество моделей с многоуровневой категоризацией предикторов заключалось в возможности объяснения и клинической интерпретации результатов прогнозирования ВГЛ.

Заключение. Многоуровневая категоризация предикторов является перспективным инструментом для объяснения прогнозных оценок в клинической медицине.

Ключевые слова: прогностические модели; многоуровневая категоризация; дихотомизация; инфаркт миокарда с подъемом сегмента ST; внутригоспитальная летальность; метод аддитивного объяснения Шепли.

Для цитирования: Шахгельдян К.И., Гельцер Б.И., Куксин Н.С., Домжалов И.Г. Многоуровневая категоризация непрерывных переменных в задачах объяснения прогнозных оценок моделей машинного обучения в клинической медицине. Врач и информационные технологии. 2023; 3: 44-57. doi: 10.25881/18110193_2023_3_44.

SHAKHGELDYAN K.I.,

DSc, Associate Professor, Vladivostok State University, Vladivostok, Russia,
e-mail: carinashakh@gmail.com

GELTSER B.I.,

Corr. Member of the RAS, DSc, Professor, Far Eastern Federal University,
Vladivostok, Russia, e-mail: boris.geltser@vvsu.ru

KUKSIN N.S.,

Far Eastern Federal University, Vladivostok, Russia, e-mail: nikita.kuksin@vvsu.ru

DOMZHALOV I.G.,

Far Eastern Federal University, Vladivostok, Russia, e-mail: igor@domzhalov.ru

MULTILEVEL CATEGORIZATION OF CONTINUOUS VARIABLES IN THE TASKS OF EXPLAINING PREDICTIVE ESTIMATES OF MACHINE LEARNING MODELS IN CLINICAL MEDICINE

DOI: 10.25881/18110193_2023_3_44

Abstract. *Aim: Comparative assessment of the quality of predictive models of in-hospital mortality (IHM) in patients with ST-segment elevation myocardial infarction (STEMI) after percutaneous coronary artery intervention (PCI), developed on the basis of predictors in continuous, dichotomous and multilevel categorical forms.*

Materials and methods: This was a single-center retrospective study analyzing data from 4677 medical records of patients with STEMI PCI who were treated at the Regional Vascular Center of Vladivostok. Two groups of patients were identified: the first consisted of 318 (6.8%) patients who died in hospital, the second — 4359 (93.2%) patients with a favorable treatment outcome. Predictive models of IHF with continuous variables were developed using multivariate logistic regression, random forest, and stochastic gradient boosting. Dichotomization of predictors was performed using grid search methods for optimal cutoff points, centroid calculation, and Shapley additive explanation (SHAP). It was proposed for multi-level categorization to use a combination of threshold values identified during dichotomization, as well as ranking cut-off thresholds using multivariate logistic regression weighting coefficients.

Results: Based on the results of a multistage analysis of indicators of the clinical and functional status of STEMI patients, new predictors of IHM were identified and validated, their categorization was performed, and prognostic models with continuous, dichotomous and multilevel categorical variables were developed (AUC: 0.885-0.902). Models whose predictors were identified using the multimetric categorization method were not inferior in accuracy to models with continuous variables and had higher quality metrics than algorithms with dichotomous predictors. The advantage of models with multilevel categorization of predictors was the ability to explain and clinically interpret the results of IHM prediction.

Conclusions: Multilevel categorization of predictors is a promising tool for explaining predictive scores in clinical medicine.

Keywords: *Predictive Models, Multi-level Categorization, Dichotomization, ST Elevation Myocardial Infarction, In-Hospital Mortality, Shapley Additive Explanation.*

For citation: *Shakhgeldyan K.I., Geltser B.I.,Kuksin N.S., Domzhalov I.G. Multilevel categorization of continuous variables in the tasks of explaining predictive estimates of machine learning models in clinical medicine. Medical doctor and information technology. 2023; 3: 44-57. doi: 10.25881/18110193_2023_3_44.*

ВВЕДЕНИЕ

В последние годы в клиническую практику все шире внедряются системы поддержки принятия врачебных решений, одной из задач которых является оценка вероятности развития неблагоприятных событий. Для реализации этой задачи разрабатываются прогностические модели на основе методов машинного обучения (МО), где в качестве предикторов используются показатели клинико-функционального статуса больных, представленные в непрерывной или категориальной формах. Категоризация определяется как преобразование непрерывных переменных с выделением не менее 2-х групп показателей, значения которых выше или ниже порогов отсечения [1]. В клинической медицине использование категориальных переменных имеет ряд преимуществ, так как отклонение анализируемых признаков от пороговых значений позволяет отнести их к факторам риска (ФР) неблагоприятных событий и дать клиническое обоснование предсказанной вероятности [2–3]. Кроме того, на основе комбинации ФР реализуется возможность формирования их фенотипов, характеризующих особенности комплексного воздействия различных признаков на результирующую переменную [4]. Вместе с тем, по мнению ряда авторов, показатели в категориальной и, в частности, дихотомической форме могут приводить к потере информации, искажению результатов анализа [5–9], увеличению доли ложноположительных [10] и ложноотрицательных [11] заключений. Альтернативой дихотомизации является подход, основанный на многоуровневой категоризации медицинских данных, позволяющий более точно классифицировать ФР и принимать оптимальные решения по их устранению [12]. Несмотря на предостережения, связанные с проблемами категоризации данных, рекомендации STROBE (Strengthening the Reporting of Observational studies in Epidemiology) подтверждают целесообразность ее использования при условии указания авторами методов, которыми она проводится [13]. При этом вопрос о необходимости выполнения процедур категоризации до настоящего времени остается в зоне активной дискуссии.

Ишемическая болезнь сердца (ИБС) занимает лидирующую позицию в структуре смертности населения от сердечно-сосудистых заболеваний

[14]. К наиболее опасным клиническим вариантам ИБС относят инфаркт миокарда с подъемом сегмента ST на электрокардиограмме (ИМпST). Одним из эффективных методов лечения ИМпST является реваскуляризация миокарда посредством чрескожного коронарного вмешательства (ЧКВ) [15]. Несмотря на совершенствование технологий ЧКВ внутригоспитальная летальность (ВГЛ) после ее выполнения по экстренным показаниям остается высокой и варьирует от 4 до 7%, что актуализирует необходимость прогнозирования неблагоприятных событий [16]. К наиболее востребованным инструментам риск-метрии относят шкалу GRACE (Global Registry of Acute Coronary Events), совершенствование которой является целью ряда исследований последних лет [17–20]. В большинстве случаев для разработки прогностических моделей базовые факторы данной шкалы дополняют новыми предикторами. Важной задачей при этом является корректное выделение их пороговых значений, что может повысить точность и объяснимость результатов прогнозирования.

Цель исследования состояла в сравнительной оценке качества прогностических моделей ВГЛ у больных ИМпST после ЧКВ, разработанных на основе предикторов в непрерывной, дихотомической и многоуровневой категориальной формах.

МАТЕРИАЛЫ И МЕТОДЫ

Данные

Проведено одноцентровое когортное ретроспективное исследование, в рамках которого анализировали данные историй болезни 4677 пациентов (3203 мужчины и 1574 женщины) с ИМпST в возрасте от 26 до 93 лет с медианой 63 года и 95% доверительным интервалом (ДИ) [62,63], поступивших в Региональный сосудистый центр «Приморская краевая клиническая больница № 1» г. Владивостока в период с 2015 по 2021 гг. [21]. Всем больным выполнено экстренное ЧКВ. Пациенты были разделены на 2 группы, первую из них составили 318 (6,8%) человек, умерших в стационаре, вторую — 4359 (93,2%) — с благоприятным исходом. Исследование соответствовало принципам Хельсинкской декларации и одобрено локальным этическим комитетом ДВФУ (протокол №8 от 8.06.2023 г.).

Критерии включения в исследование: больные с подтвержденным ИМпST и выполнением ЧКВ в первые сутки стационарного лечения. Критерии невключения: нестабильная стенокардия, ИМ без подъема сегмента ST и отсутствие показаний для ЧКВ.

Клинико-функциональный статус больных оценивали в первые сутки стационарного лечения с помощью 136 факторов, наиболее значимые из которых представлены в таблице 1. Данные включали 5 признаков шкалы GRACE: возраст больных, класс острой сердечной недостаточности (ОСН) по Т. Killip, частоту сердечных сокращений (ЧСС), систолическое артериальное давление (САД), концентрацию креатинина (Cr) в сыворотке крови. Анализировали содержание в крови нейтрофилов (NEUT), эозинофилов (EOS), уровень тромбокриты (PCT), глюкозы (Gl), эхокардиографические индикаторы, в том числе фракцию выброса (ФВ) левого желудочка (ЛЖ) и др.

Конечная точка исследования была представлена показателем ВГЛ у больных ИМпST после ЧКВ от всех причин в форме категориального бинарного признака («отсутствие» или «развитие»).

Методы статистического анализа и машинного обучения. Распределение данных по тесту Колмогорова-Смирнова отличалось от нормального, в связи с чем использовали непараметрические методы статистики. Показатели были представлены медианой (Me) и межквартильными интервалами (Q1; Q3), для межгрупповых сравнений непрерывных переменных применяли тест Манна-Уитни, а для категориальных — χ^2 . Отношение шансов (ОШ) и их 95% ДИ рассчитывали точным тестом Фишера. Статистически значимыми считали различия при $p < 0,01$.

Прогностические модели ВГЛ с предикторами в непрерывной форме разрабатывали методами многофакторной логистической регрессии (МЛР), случайного леса (СЛ) и стохастического градиентного бустинга (СГБ). Модели, в которых использовались категориальные предикторы, разрабатывали только на основе МЛР и СГБ. Их качество оценивали по 6 метрикам: площадь под ROC-кривой (AUC), чувствительность (Sen), специфичность (Spec), F1-score, positive predictive value (PPV), negative predictive value (NPV).

Для дихотомизации потенциальных предикторов использовали методы оптимизации на сетке с шагом $\Delta = (\max - \min) / 100$: минимизацию

p-value — Min(p-value), максимизацию AUC — Max(AUC), квартили, метод центроидов и аддитивного объяснения Шепли (SHAP) [22, 23]. Последний применяли также для оценки степени влияния предикторов модели на конечную точку. Для выполнения многоуровневой категоризации нами предложено 2 новых подхода. Первый из них предполагал использование пороговых значений, полученных по результатам анализа SHAP-value (рис. 1) Второй, мультиметрический, объединял все пороговые значения, выделенные при дихотомизации показателей различными методами, в том числе многоуровневой категоризации по SHAP. При этом близкие пороговые значения объединились в одно путем усреднения. Кроме того, в работе применяли 2 процедуры категоризации, предложенные ранее. В первой использовали медианы анализируемых факторов в группах сравнения и равноудаленные от них значения (центроиды), с помощью которых выделяли 4 категории для каждого показателя [22]. Вторая процедура была связана с выделением 4-х категорий для каждой переменной по результатам оценки их медиан, 2 и 3 квартили [24].

Датасет был разделен на 2 выборки: для обучения и кросс-валидации (80%) и для итогового тестирования (20%). Процедура обучения и кросс-валидации выполнялась методом стратифицированного k-Folders на 10 выборках. Усредненные метрики качества AUC, Sen и Spec использовали для выбора лучшей модели, отбора и валидации предикторов и подбора оптимальных гиперпараметров путем перебора по сетке допустимых значений. Для заключительного тестирования лучшие модели МЛР, СЛ и СГБ с оптимальными параметрами и гиперпараметрами обучали на 80% датасета, а проверяли на подгруппе для итогового тестирования. Для доверительной оценки метрик качества процедуру повторяли 500 раз, выполняя первоначальное разделение случайным образом методом Монте-Карло.

Анализ данных и построение моделей проводились в Python с открытым исходным кодом, версии 3.9.16.

ДИЗАЙН ИССЛЕДОВАНИЯ

Дизайн исследования включал 5 этапов. На первом из них с помощью тестов межгрупповых

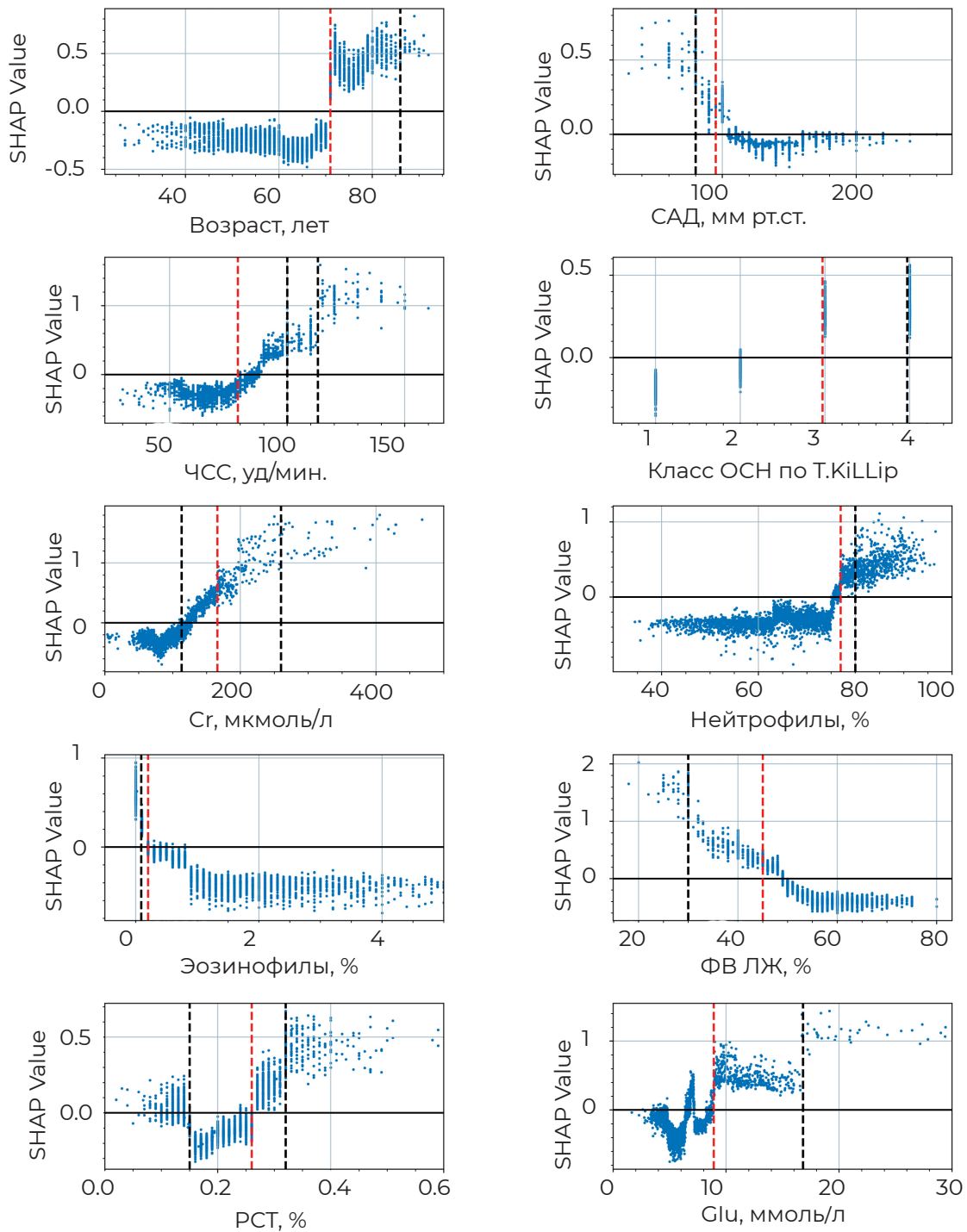


Рисунок 1 — Оценка влияния непрерывных показателей и их пороговых значений на конечную точку по методу SHAP. Примечание: пунктирные линии обозначают пороги отсечения. Сокращения: ФВ ЛЖ — фракция выброса левого желудочка, Glu — глюкоза, РСТ — тромбоциты; САД — систолическое артериальное давление, ЧСС — частота сердечных сокращений, Cr — креатинин.

сравнений формировался пул потенциальных предикторов ВГЛ (табл. 1). На втором этапе исследования с помощью методов МО разрабатывали прогностические модели ВГЛ с предикторами в непрерывной форме, включающие 5 базовых факторов шкалы GRACE. Для повышения точности прогноза структуру моделей пошагово дополняли новыми предикторами,

выделенными на первом этапе исследования, при условии статистически значимого различия на уровне $p\text{-value} < 0.01$. Прогностическая значимость предиктора подтверждалась, если после его включения в модель значение AUC возрастало. На третьем этапе с помощью различных методов выделения пороговых значений осуществляли бинаризацию непрерывных переменных

Таблица 1 — Клинико-функциональная характеристика больных ИМпST

Предиктор	Группа 1 (n = 318)	Группа 2 (n = 4357)	ОШ (95%) ДИ	p-value
Пол: Ж, абс. (%)	142 (44.65)	1332 (30.5)	1.8 [1.5; 2.3]	<0.000001
Возраст, лет	71 (63; 78)	62 (55; 69)	–	<0.000001
Рост, см	168 (164; 174)	170 (165; 176)	–	0.000001
Вес, кг	78 (70; 85)	80 (71; 90)	–	0.000012
САД, мм рт.ст.	110 (90; 130)	130 (120; 150)	–	<0.000001
ДАД, мм рт.ст.	70 (60; 80)	80 (75; 90)	–	<0.000001
ЧСС, уд/мин	86 (72; 100)	72 (65; 80)	–	<0.000001
Сг, мкмоль/л	130 (96; 193.3)	97 (81; 114.8)	–	<0.000001
Класс ОЧН по Т.Killip				
I	71 (22.33%)	2726 (62.6%)	0.17 [0.13; 0.23]	<0.000001
II	58 (18.2%)	867 (19.9%)	0.9 [0.67; 1.20]	0.508052
III	66 (20.75)	479 (11)	2.1 [1.6; 2.8]	<0.000001
IV	123 (38.7)	269(6.18)	9.6 [7.4; 12.4]	<0.000001
III-IV	189 (59.4)	748 (17.2)	7.1 [5.6; 9]	<0.000001
ФВ ЛЖ, %	46.5 (38; 54.8)	56 (50; 61)	–	<0.000001
NEUT, %	81.3 (75.75; 86.5)	66.7 (59.1; 74.9)	–	<0.0001
EOS, %	0.1 (0.00; 0.3)	0.9 (0.3; 1.9)	–	<0.000001
Glu, ммоль/л	7.9 (6.3; 10.31)	5.8 (5.1; 7)	–	<0.000001
PCT, %	0.22 (0.17; 0.28)	0.2 (0.16; 0.24)	–	0.0012
Передний ИМ, абс. (%)	177 (55.66%)	2017 (46.27%)	1.5 [1.16; 1.83]	0.00147
ФП, абс. (%)	129 (40.57%)	772 (17.69%)	3.2 [2.51; 4.02]	<0.000001
СД 2, абс. (%)	99 (31.13%)	831 (19.05%)	1.9 [1.50; 2.46]	<0.000001
ХБП, абс. (%)	83 (26.1%)	677 (15.5%)	1.97 [1.5; 2.6]	<0.000001

Сокращения: ФВ ЛЖ — фракция выброса левого желудочка, Glu — глюкоза, PCT — тромбокрит; САД — систолическое артериальное давление, ДАД — диастолическое артериальное давление, ЧСС — частота сердечных сокращений, Сг — креатинин, ФП — фибрилляция предсердий, СД 2 – сахарный диабет второго типа, ХБП — хроническая болезнь почек.

и на их основе разрабатывали прогностические модели ВГЛ. На четвертом этапе исследования проводили многоуровневую категоризацию переменных. В рамках каждого подхода формировали группы потенциальных предикторов ВГЛ. Для оценки степени их влияния на конечную точку были разработаны модели МЛР, весовые коэффициенты которых использовали для кодирования многоуровневых категориальных предикторов. На пятом этапе, используя СГБ, разработаны 4 новых прогностических модели ВГЛ, предикторы которых были получены разными методами многоуровневой категоризации.

РЕЗУЛЬТАТЫ

Межгрупповой анализ клинико-демографических и лабораторных показателей продемонстрировал, что большинство из них, в том числе, все предикторы шкалы GRACE имеют статистически значимые различия (табл. 1). Среди умерших преобладали лица старшего возраста, женского пола и более низкого роста. Для первой группы пациентов было характерным наличие 3 и 4 класса ОЧН по Т. Killip, меньшие значения САД, ДАД, ФВ ЛЖ, EOS и некоторых других показателей, более высокие уровни ЧСС, Cr, NEUT. У них чаще фиксировался передний ИМ, сахарный диабет 2-го типа (СД2), фибрилляция предсердий (ФП) и хроническая болезнь почек (ХБП).

На втором этапе исследования были разработаны прогностические модели ВГЛ, где помимо 5 базовых показателей шкалы GRACE в непрерывной форме использовали новые факторы. Последние были выделены в результате тестирования пула потенциальных предикторов в прогностических моделях. К ним относились показатели ФВ ЛЖ, NEUT, EOS, PCT и Glu, при последовательном включении которых в структуру моделей МЛР, СГБ и СЛ, точность прогноза по метрике AUC возрастала, что свидетельствовало о предиктивной ценности анализируемых факторов. Необходимо также отметить, что сопоставление значений AUC модели, разработанной нами ранее на этой же выборке больных с использованием только предикторов шкалы GRACE [25], и модели, дополненной новыми предикторами, демонстрировало более высокую точность последней (AUC — 0,9 vs 0,836).

Сопоставление предсказательной ценности разработанных моделей МЛР, СГБ и СЛ показало, что они обладают высокой прогностической способностью и имеют близкие значения метрик качества при кросс-валидации и итоговом тестировании (AUC варьировала от 0,884 до 0,9), что свидетельствует об отсутствии их переобучения и хороших обобщающих свойствах (табл. 2).

На третьем этапе исследования выполняли дихотомизацию предикторов ВГЛ в

Таблица 2 — Оценка точности прогностических моделей ВГЛ у больных ИМпСТ после ЧКВ с использованием предикторов в непрерывной форме

Метрики	Кросс-валидация			Итоговое тестирование		
	МЛР	СГБ	СЛ	МЛР	СГБ	СЛ
AUC	0.9 [0.885; 0.916]	0.891 [0.871; 0.911]	0.885 [0.87; 0.9]	0.9 [0.841; 0.959]	0.892 [0.834; 0.951]	0.884 [0.824; 0.944]
Sen	0.843 [0.81; 0.877]	0.825 [0.779; 0.872]	0.796 [0.749; 0.843]	0.843 [0.715; 0.972]	0.824 [0.692; 0.957]	0.798 [0.672; 0.925]
Spec	0.836 [0.824; 0.849]	0.816 [0.797; 0.835]	0.806 [0.784; 0.828]	0.838 [0.807; 0.868]	0.819 [0.783; 0.855]	0.806 [0.766; 0.846]
PPV	0.165 [0.152; 0.178]	0.1146 [0.131; 0.161]	0.136 [0.124; 0.148]	0.168 [0.141; 0.194]	0.15 [0.125; 0.175]	0.138 [0.101; 0.167]
NPV	0.993 [0.992; 0.995]	0.992 [0.99; 0.994]	0.991 [0.989; 0.993]	0.993 [0.987; 0.999]	0.992 [0.986; 0.998]	0.991 [0.985; 0.996]
F-score	0.275 [0.256; 0.295]	0.248 [0.226; 0.27]	0.231 [0.214; 0.249]	0.278 [0.239; 0.317]	0.253 [0.215; 0.291]	0.235 [0.189; 0.28]

непрерывной форме методами поиска оптимального порога отсека на сетке, методом SHAP и с помощью расчета центраида (табл. 3). Использование пороговых значений, отклонение от которых ассоциируется с увеличением вероятности ВГЛ, позволяет рассматривать бинаризованные данные как ФР неблагоприятных событий. ФР кодируется "1", если значение предиктора превышает порог с постфиксом "+" или не достигает его — с постфиксом "-", в остальных случаях — "0".

Результаты исследования показали, что пороговые значения, полученные различными методами бинаризации, часто отличались между собой. Так, например, порог отсека для показателя ФВ ЛЖ по версии SHAP составлял <45%, в то время как при максимизации AUC точка отсека фиксировалась на уровне 51% (табл. 3). При этом пороговые значения, определенные алгоритмом SHAP, были ближе всего к критериальным границам, верифицированным методом Min(p-value). Необходимо отметить, что метод SHAP позволяет не только определить пороговые границы, но и оценить интенсивность влияния на ВГЛ показателей, значения которых находятся в "зоне риска". К

категориальным факторам, выделенным этим методом, относились следующие признаки: возраст > 70 лет, ЧСС > 89 уд./мин, САД < 95 мм рт.ст., класс ОСН по Т. Killip > 2, Cr > 166 ммоль/л, ФВ ЛЖ < 45%, NEUT > 77%, EOS < 0.22%, PCT > 0.32%, Glu > 8.9 ммоль/л (рис. 1). На примере диаграммы ФВ ЛЖ видно, что вероятность ВГЛ последовательно увеличивается в диапазоне значений данного показателя 44% — 31% и резко возрастает при его уровне < 30%. Повышение концентрации глюкозы в крови более 8.9 ммоль/л увеличивает риск ВГЛ, но вероятность фатального исхода значительно возрастает при ее уровне более 17 ммоль/л. Оценка динамики изменений SHAP-value, позволяющая объяснить взаимосвязь различных значений предикторов с конечной точкой исследования, явилась основанием для использования этого метода в процедурах многоуровневой категоризации.

На основе МЛР были разработаны прогностические модели ВГЛ с дихотомическими предикторами (табл. 4). Было установлено, что большинство из них обладает высокой предиктивной способностью вне зависимости от метода определения пороговых значений (AUC: 0.885 — 0.896).

Таблица 3 — Дихотомизация непрерывных предикторов ВГЛ с помощью различных методов

Предиктор	Min(p-value)	Max(AUC)	Центроид	SHAP
Возраст	70+	65+	66.5+	71+
САД	92-	112-	120-	95-
ЧСС	95+	82+	79+	89+
Класс ОСН по Т. Killip	4	3+	3+	3+
Cr, ммоль/л	188.6+	122.9+	113.3+	166+
NEUT, %	78.8+	75+	74.0+	77+
EOS, %	0.3-	0.3-	0.5-	0.2-
ФВ ЛЖ, %	31.0-	51-	51-	45-
PCT, %	0.36+	0.22+	0.21+	0.32+
Glu, ммоль/л	8.4+	6.5+	6.9+	8.9+

Сокращения: ФВ ЛЖ — фракция выброса левого желудочка (ЛЖ), Glu — глюкоза, PCT — тромбокрит; САД — систолическое артериальное давление, ЧСС — частота сердечных сокращений, Cr — креатинин.

Таблица 4 — Оценка точности прогностических моделей ВГЛ на основе предикторов в дихотомической форме

Метрики	SHAP	Min(p-value)	Max(AUC)	Центроиды
AUC	0.885 [0.882; 0.887]	0.876 [0.873; 0.879]	0.896 [0.893; 0.898]	0.888 [0.885;0.890]
Sen	0.815 [0.809; 0.821]	0.815 [0.809; 0.821]	0.815 [0.809; 0.821]	0.815 [0.808;0.821]
Spec	0.825 [0.823; 0.826]	0.826 [0.824; 0.828]	0.821 [0.819; 0.822]	0.823 [0.822;0.825]
PPV	0.149 [0.148; 0.15]	0.147 [0.116; 0.178]	0.148 [0.147; 0.149]	0.15 [0.149; 0.151]
NPV	0.991 [0.991;0.992]	0.991 [0.991;0.992]	0.991 [0.991;0.992]	0.991 [0.991;0.992]
F-score	0.25 [0.248; 0.252]	0.251 [0.25; 0.253]	0.251 [0.25; 0.253]	0.254 [0.252;0.256]

На четвертом этапе исследования на основе использования 4-х методов многоуровневой категоризации были сформированы группы ФР ВГЛ. Первый пул ФР был получен при анализе SHAP-value (рис.). Второй — расширял первый пул ФР дихотомическими пороговыми значениями, полученными на предыдущем этапе. Третья группа включала медианы предикторов в группах сравнения и их центроиды, а четвертая — использовала пороговые значения, соответствующие квартилям предикторов. Для кодирования значений многоуровневых категориальных предикторов использовали весовые коэффициенты (ВК) моделей МЛР, разработанных для каждой группы ФР (табл. 5).

На пятом этапе на основе многоуровневых предикторов, полученных различными методами, были разработаны 4 прогностические модели ВГЛ на основе СГБ (табл. 6). Лучшие прогностические свойства демонстрировала модель с предикторами, выделенными методом мультиметрической категоризации. Последняя была сопоставима по точности с моделью, включающей непрерывные переменные, но в отличие от нее позволяла объяснять прогноз ВГЛ по результатам оценки пороговых значений и ВК ее предикторов (табл. 5). С учетом этих данных было установлено, что у больных ИМпСТ после ЧКВ наибольшая вероятность развития неблагоприятного исхода ассоциируется с показателями ФВ ЛЖ < 31% (ВК — 1.735) и Cr > 189 ммоль/л (ВК — 1.489). К менее значимым факторам риска ВГЛ относятся ЧСС > 95 уд/мин (ВК — 1.151), уровень глюкозы > 17 ммоль/л (ВК — 1.12) и содержание эозинофилов < 0.09% (ВК — 0.935).

Интенсивность воздействия на конечную точку 4-го класса ОЧН по T. Killip, возраста > 71 года, содержания нейтрофилов > 77% и концентрации креатинина в диапазоне 166–189 ммоль/л характеризовалась значениями ВК на уровне 0.75-0.86. Остальные категориальные значения ФР оказывали на ВГЛ существенно меньшее влияние. Сравнительная оценка точности прогностических моделей с предикторами, выделенными методами дихотомизации и многоуровневой категоризации, демонстрировала преимущества последних, что подтверждалось статистически значимыми различиями метрики AUC (p-value<0,00001).

ОБСУЖДЕНИЕ

В последние годы на основе методов МО разрабатываются прогностические модели, применение которых в клинической практике ограничивается сложностью интерпретации результатов прогноза. Перспективными инструментами для решения этой задачи являются алгоритмы объяснимого искусственного интеллекта (ОИИ), к которым относят определение пороговых значений, используемых в моделях ФР, и их ранжирование по интенсивности влияния на реализацию конечной точки [26]. Одним из подходов к определению пороговых значений ФР ВГЛ является их категоризация, которая позволяет детализировать соотношения показателей клиничко-функционального статуса больных ИМпСТ с результирующей переменной. По данным литературы, наиболее доступным методом многоуровневой категоризации является дескриптивная статистика с расчетом медиан,

Таблица 5 — Весовые коэффициенты и пороги предикторов, полученные методами многоуровневой категоризации

Предикторы	SHAP		Мультиметрическая категоризация		Медианы групп и центроид		Квартили	
	Пороги	ВК	Пороги	ВК	Пороги	ВК	Пороги	ВК
Возраст	[71;86] 86+	0.902 0.839	[65;71] 71+	0.112 0.873	[66.5;71] 71+	0.67 0.857	[63;70] 70+	0.018 0.85
САД	80- [80;95]	0.975 0.09	60- [60;92] [92;112] [112; 120]	0.276 0.54 0.352 0.364	110- [110;120]	0.842 0.146	120-	0.367
ЧСС	[79;100] [100;113] 113+	0.422 0.86 1.343	[82;95] 95+	0.538 1.151	86+	1.012	81+	0.976
класс ОСН по Т. Killip	3 4	0.502 0.879	3 4	0.512 0.859	3 4	0.261 0.689	3 4	0.403 1.025
Cr	[113;166] [166;260] 260+	0.431 1.094 1.694	[123;166] [166;189] 189+	0.456 0.774 1.489	[97;113] [113;130] 130+	0.239 0.434 1.062	[98;117] 117+	0.263 1.035
Нейтрофилы	[77;80] 80+	0.847 0.782	[75;77] [77;79] 79+	0.405 0.78 0.75	[74;81.3] 81.3+	0.5 0.758	75.8+	0.832
Эозинофилы	0.09- [0.09;0.2]	0.883 0.484	0.09- [0.09;0.3]	0.935 0.371	0.1- [0.1;0.5] [0.5;0.9]	1.07 0.566 0.695	0.2- [0.2;0.8] [0.8;1.8]	0.913 0.41 0.167
ФВ ЛЖ	30- [30;45]	1.255 0.794	31- [31;45] [45;51]	1.735 0.566 0.311	46.5- [46.5;51]	0.916 0.422	49- [49;56]	0.815 0.039
PCT	0.15- [0.15;0.26] [0.26;0.32] 0.32+	0.304 0 0.374 0.658	[0.31;0.36] 0.36+	0.551 0.582	[0.2;0.21] 0.22+	0.324 0.258	0.16- 0.25+	0.56 0.565
Glu	[8.9;17] 17+	0.856 1.373	[6.5;6.9] [6.9;8.9] [8.9;17] 17+	0.368 0 0.648 1.12	[5.8;6.9] [6.9;7.9] 7.92+	0.077 0.236 0.949	[5.83;7.1] 7.1+	0.33 0.77

Сокращения: ФВ ЛЖ — фракция выброса левого желудочка (ЛЖ), Glu — глюкоза, PCT — тромбокрит; САД — систолическое артериальное давление, ЧСС — частота сердечных сокращений, Cr — креатинин, ВК — весовые коэффициенты.

квартилей или квантилей [1, 9, 27]. Вместе с тем большая часть критических замечаний по категоризации связана именно с этим подходом, что обусловлено, прежде всего, зависимостью таких пороговых значений от конкретной выборки, отсутствием взаимосвязи с клиническим

контекстом, игнорированием возможных нелинейных отношений и др. Другой подход основан на выделении порогового значения известного из практики как выход за пределы нормы, который не связан с клиническим контекстом прогнозируемого неблагоприятного события [27].

Таблица 6 — Оценка точности прогностических моделей ВГЛ на основе предикторов с многоуровневой категоризацией

Метрики	Мультиметрическая категоризация	Категоризация методом SHAP	Медианы групп и центроид	Квартили
AUC	0.902 [0.9; 0.903]	0.897 [0.894; 0.898]	0.894 [0.892; 0.896]	0.887 [0.885; 0.889]
Sen	0.825 [0.819; 0.831]	0.825 [0.82; 0.83]	0.825 [0.819; 0.831]	0.8 [0.795; 0.805]
Spec	0.834 [0.833; 0.836]	0.83 [0.828; 0.831]	0.817 [0.815; 0.818]	0.806 [0.805; 0.808]
PPV	0.157 [0.156; 0.158]	0.153 [0.152; 0.154]	0.146 [0.145; 0.147]	0.14 [0.139; 0.141]
NPV	0.991 [0.991; 0.991]	0.992 [0.992; 0.992]	0.991 [0.991; 0.991]	0.991 [0.991; 0.991]
F-score	0.259 [0.258; 0.261]	0.258 [0.256; 0.259]	0.249 [0.248; 0.25]	0.24 [0.239; 0.242]

Альтернативным методом, учитывающим клинический контекст, является поиск оптимальных пороговых значений на основе минимизации или максимизации целевых функций, например, $\text{Min}(p\text{-value})$ или $\text{Max}(AUC)$. По нашему мнению, категоризация переменных должна рассматриваться только в рамках решения конкретной клинической задачи, хотя выделенные предикторы могут представлять интерес для реализации других прогностических исследований.

В нашей работе на примере базы данных больных ИМПСТ после ЧКВ анализировали эффективность различных методов поиска пороговых значений предикторов, отклонение от которых повышало их предиктивный потенциал и позволяло отнести к ФР ВГЛ. Было установлено, что метод SHAP, который рассматривают как технологию ОИИ, является полезным инструментом категоризации за счет эффективного определения порогов отсечения, в том числе, для многоуровневой категоризации и анализа взаимосвязи предикторов в непрерывной и категориальной формах с конечной точкой исследования. Так, например, с помощью данного метода для показателя ФВ ЛЖ наряду с границей 45% был выделен порог на уровне 31%, ниже которого риск ВГЛ резко возрастает. В то же время было показано, что многоуровневые категориальные предикторы, полученные путем объединения данных SHAP с результатами других методов дихотомизации, обеспечивают более высокую прогностическую точность.

В нашем исследовании потенциальные риски потери информации при

использовании новых методов категоризации были преодолены за счет детализации знаний о взаимосвязи отдельных ФР или их комбинаций с конечной точкой. Это подтверждалось сопоставлением критериев качества прогностических моделей с предикторами в непрерывной и многоуровневой категориальной формах. Так, для лучшей модели с непрерывными предикторами AUC составляла 0.9, а при использовании мультиметрической категоризации — 0.902. Методы категоризации, основанные на дескриптивных статистиках, не компенсировали потерю прогностической точности за счет привнесения новой информации.

Ограничения исследования связаны с его ретроспективным характером, необходимостью валидации моделей на данных из других лечебных учреждений.

ЗАКЛЮЧЕНИЕ

В настоящем исследовании на основе базы данных пациентов с ИМПСТ после ЧКВ апробированы 2 новых подхода к многоуровневой категоризации предикторов ВГЛ: мультиметрическая категоризация и метод SHAP. С помощью разработанных прогностических моделей ВГЛ показано, что предложенные авторами процедуры категоризации обеспечивают высокую точность и прозрачность результатов прогнозирования.

Источник финансирования. Исследование выполнено в рамках проекта Российского научного фонда (РНФ) № 23-21-00250, <https://rscf.ru/project/23-21-00250/>

ЛИТЕРАТУРА/REFERENCES

1. Mabikwa OV, Greenwood DC, Baxter PD, Fleming SJ. Assessing the reporting of categorised quantitative variables in observational epidemiological studies. *BMC Health Serv Res.* 2017; 17(1): 201. doi:10.1186/s12913-017-2137-z.
2. MacCallum RC, Zhang S, Preacher KJ, Rucker DD. On the practice of dichotomization of quantitative variables. *Psychol Methods.* 2002; 7(1): 19-40. doi:10.1037/1082-989x.7.1.19.
3. Gupta R, Day CN, Tobin WO, Crowson CS. Understanding the effect of categorization of a continuous predictor with application to neuro-oncology. *Neurooncol Pract.* 2021; 9(2): 87-90. doi:10.1093/nop/npab049.
4. Гельцер Б.И., Шахгельдян К.И., Рублев В.Ю. и др. Фенотипирование факторов риска и прогнозирование внутригоспитальной летальности у больных ишемической болезнью сердца после коронарного шунтирования на основе методов объяснимого искусственного интеллекта // Российский кардиологический журнал. — 2023. — Т.28. — №4. — С.85-93. [Geltser BI, Shakhgeldyan KI, Rublev VYu, Domzhalov IG, Tsivanyuk MM, Shekunova OI. Phenotyping of risk factors and prediction of inhospital mortality in patients with coronary artery disease after coronary artery bypass grafting based on explainable artificial intelligence methods. *Russian Journal of Cardiology.* 2023; 28(4): 5302. (In Russ.) doi:10.15829/1560-4071-2023-5302.
5. Altman DG, Lausen B, Sauerbrei W, Schumacher M. Dangers of using «optimal» cutpoints in the evaluation of prognostic factors. *J Natl Cancer Inst.* 1994; 86(11): 829-835. doi:10.1093/jnci/86.11.829.
6. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *BMC Med.* 2015; 13: 1. doi:10.1186/s12916-014-0241-z.
7. Dawson NV, Weiss R. Dichotomizing continuous variables in statistical analysis: a practice to avoid. *Med Decis Making.* 2012; 32(2): 225-226. doi:10.1177/0272989X12437605.
8. Salis Z, Gallego B, Sainsbury A. Researchers in rheumatology should avoid categorization of continuous predictor variables. *BMC Med Res Methodol.* 2023; 23(1): 104. doi:10.1186/s12874-023-01926-4.
9. Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ.* 2006; 332(7549): 1080. doi:10.1136/bmj.332.7549.1080.
10. Austin PC, Brunner LJ. Inflation of the type I error rate when a continuous confounding variable is categorized in logistic regression analyses. *Stat Med.* 2004; 23(7): 1159-1178. doi:10.1002/sim.1687.
11. Streiner DL. Breaking up is hard to do: the heartbreak of dichotomizing continuous data. *Can J Psychiatry.* 2002; 47(3): 262-266. doi: 10.1177/070674370204700307.
12. Chen Y, Huang J, He X, et al. A novel approach to determine two optimal cut-points of a continuous predictor with a U-shaped relationship to hazard ratio in survival data: simulation and application. *BMC Med Res Methodol.* 2019; 19(1): 96. Published 2019. doi:10.1186/s12874-019-0738-4.
13. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet.* 2007; 370(9596): 1453-1457. doi:10.1016/S0140-6736(07)61602-X.
14. The World Health Organization, The top 10 causes of death [Internet]. 2019. Available from: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death> [cited 2023 Nov 30].
15. Ibáñez B, James S, Agewall S, et al. 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation. *Rev Esp Cardiol (Engl Ed).* 2017; 70(12): 1082. doi:10.1016/j.rec.2017.11.010.

16. Pfuntner A, Wier LM, Stocks C. Most Frequent Procedures Performed in U.S. Hospitals, 2011. In: Healthcare Cost and Utilization Project (HCUP) Statistical Briefs. Rockville (MD): Agency for Healthcare Research and Quality (US); October 2013.
17. Wang JJ, Fan Y, Zhu Y, et al. Biomarkers enhance the long-term predictive ability of the KAMIR risk score in Chinese patients with ST-elevation myocardial infarction. *Chin Med J*. 2019; 132(1): 30-41. doi:10.1097/CM9.000000000000015.
18. Liu XJ, Wan ZF, Zhao N, et al. Adjustment of the GRACE score by HemoglobinA1c enables a more accurate prediction of long-term major adverse cardiac events in acute coronary syndrome without diabetes undergoing percutaneous coronary intervention. *Cardiovasc Diabetol*. 2015; 14: 110. doi: 10.1186/s12933-015-0274-4.
19. Chen X, Shao M, Zhang T, et al. Prognostic value of the combination of GRACE risk score and mean platelet volume to lymphocyte count ratio in patients with ST-segment elevation myocardial infarction after percutaneous coronary intervention. *Exp Ther Med*. 2020; 19(6): 3664-3674. doi: 10.3892/etm.2020.8626.
20. Wenzl FA, Kraler S, Ambler G, et al. Sex-specific evaluation and redevelopment of the GRACE score in non-ST-segment elevation acute coronary syndromes in populations from the UK and Switzerland: a multinational analysis with external cohort validation. *Lancet*. 2022; 400(10354): 744-756. doi: 10.1016/S0140-6736(22)01483-0.
21. Гельцер Б.И., Шахгельдян К.И., Домжалов И.Г., и др. Прогностическая оценка клинико-функционального статуса пациентов с инфарктом миокарда с подъемом сегмента ST после чрескожного коронарного вмешательства. Свидетельство о регистрации базы данных 2023622740, 10.08.2023. Заявка № 2023622516 от 28.07.2023. [Geltser BI, Shakhgelyan KI, Domzhalov IG, et al. Prognosticheskaya ocenka kliniko-funkcional'nogo statusa pacientov s infarktomiokarda s pod'emom segmenta ST posle chreskozhnogo koronarnogo vmeshatel'stva. Certificate of registration of the database 2023622740, 10.08.2023. Request № 2023622516. 28.07.2023.]

22. Valente F, Henriques J, Paredes S, et al. A new approach for interpretability and reliability in clinical risk prediction: Acute coronary syndrome scenario. *Artif Intell Med.* 2021; 117: 102113. doi: 10.1016/j.artmed.2021.102113.
23. Lundberg SM, Lee SI. A unified approach to interpreting model predictions, in: *Advances in Neural Information Processing Systems. Proceedings of the 31st Annual Conference on Neural Information Processing Systems; 2017 Dec 04-09; Long Beach, USA.* doi: 10.48550/arXiv.1705.07874.
24. Evenson KR, Wen F, Herring AH. Associations of Accelerometry-Assessed and Self-Reported Physical Activity and Sedentary Behavior With All-Cause and Cardiovascular Mortality Among US Adults. *Am J Epidemiol.* 2016; 184(9): 621-632. doi: 10.1093/aje/kww070.
25. Гельцер Б.И., Шахгельдян К.И., Домжалов И.Г. и др. Прогнозирование внутригоспитальной летальности у больных острым инфарктом миокарда с подъемом сегмента ST после чрескожного коронарного вмешательства // *Российский кардиологический журнал.* — 2023. — Т.28 — №6. — С.31-39. [Geltser BI, Shahgeldyan KI, Domzhalov IG, et al. Prediction of in-hospital mortality in patients with ST-segment elevation acute myocardial infarction after percutaneous coronary intervention. *Russian Journal of Cardiology.* 2023; 28(6): 5414. (In Russ.)] doi: 10.15829/1560-4071-2023-5414.
26. Molnar C. *Interpretable Machine Learning. A Guide for Making Black Box Models Explainable* [Internet]. North Charleston: Independently published. 2023. [cited 2023 Nov 30]. Available from <https://christophm.github.io/interpretable-ml-book>.
27. Turner EL, Dobson JE, Pocock SJ. Categorisation of continuous risk factors in epidemiological publications: a survey of current practice. *Epidemiol Perspect Innov.* 2010; 7: 9. doi:10.1186/1742-5573-7-9.

ЖУРАВЛЕВ Д.М.,

д.э.н., Научно-исследовательский институт Социальных Систем при МГУ имени М.В. Ломоносова, г. Москва, Россия, e-mail: jdenis@niiss.ru

КОПЫЛОВ Ф.Ю.,

д.м.н., Первый Московский государственный медицинский университет им. И.М. Сеченова, г. Москва, Россия, e-mail: kopylov_f_yu@staff.sechenov.ru

ЧААДАЕВ В.К.,

д.э.н., Научно-исследовательский институт Социальных Систем при МГУ имени М.В. Ломоносова, г. Москва, Россия, e-mail: vkchaadaev@niiss.ru

АРДАТОВ С.В.,

к.м.н., Самарский государственный медицинский университет, г. Самара, Россия, e-mail: ardatov67@mail.ru

ЧААДАЕВ К.В.,

Инновационный Технологический Центр МОЛНЕТ, г. Москва, Россия; Московский государственный технический университет имени Н.Э. Баумана, г. Москва, Россия, e-mail: kchaadaev@molnet.ru

АВТОМАТИЗИРОВАННЫЙ КОМПЛЕКС МУЛЬТИДИСЦИПЛИНАРНОЙ НЕЙРОСЕТЕВОЙ ПОДДЕРЖКИ ВРАЧЕБНЫХ РЕШЕНИЙ ПРИ ЛЕЧЕНИИ ИШЕМИЧЕСКОЙ БОЛЕЗНИ СЕРДЦА

DOI: 10.25881/18110193_2023_3_58

Аннотация.

В статье рассмотрены методы и процедуры разработки нейросетевой системы поддержки принятия решения при выборе тактики оперативного вмешательства на коронарных сосудах сердца, предназначенной для консультирования широкого круга практикующих кардиологов и кардиохирургов при принятии решения о тактике проведения оперативного вмешательства при состояниях пациентов, связанных с нарушением проводимости коронарных сосудов. Основываясь на математической модели, учитывающей ряд факторов и опыт исходов предыдущих операций, нейросетевая система предлагает выбор между аортокоронарным шунтированием и чрескожным коронарным вмешательством. Определённое системой решение может служить дополнительным голосом и фактором для окончательного принятия коллегиального решения в сложных клинических случаях. Правильно принятое решение влияет на сроки восстановления пациента после операции, качество жизни после восстановления, возможность продолжать трудовую деятельность после лечения. Нейросетевая система поддержки принятия решения в области кардиохирургии выполнена в виде стандартного приложения для персонального компьютера со специфическими техническими характеристиками, позволяющими обрабатывать большой массив данных. Доступ к системе может получить любой врач кардиолог или кардиохирург, зарегистрированный в системе и прошедший валидацию. Созданный комплекс призван обеспечить учреждения системы здравоохранения цифровым продуктом и сервисом отечественного производства на основе нового технологического уклада. Работа выполнена при поддержке Фонда содействия развитию малых форм предприятий в научно-технической сфере (Фонд содействия инновациям).

Ключевые слова: диагностика и лечение, ишемическая болезнь сердца, математическая модель, нейронная сеть, программное обеспечение, сердечно-сосудистые заболевания.

Для цитирования: Журавлев Д.М., Копылов Ф.Ю., Чаадаев В.К., Ардатов С.В., Чаадаев К.В. Автоматизированный комплекс мультидисциплинарной нейросетевой поддержки врачебных решений при лечении ишемической болезни сердца. *Врач и информационные технологии*. 2023; 3: 58-71. doi: 10.25881/18110193_2023_3_58.

ZHURAVLEV D.M.,

DSc, Institute of Social Systems, Moscow, Russia,
e-mail: jdenis@niiss.ru

KOPYLOV F.YU.,

DSc, I.M. Sechenov First Moscow State Medical University, Moscow, Russia,
e-mail: kopylov_f_yu@staff.sechenov.ru

CHAADAEV V.K.,

DSc, Institute of Social Systems, Moscow, Russia,
e-mail: vkchaadaev@niiss.ru

ARDATOV S.V.,

PhD, Samara State Medical University, Samara, Russia,
e-mail: ardatov67@mail.ru

CHAADAEV K.V.,

ITC MOLNET, Moscow, Russia; Bauman Moscow State Technical University, Moscow, Russia,
e-mail: kchaadaev@molnet.ru

AUTOMATED COMPLEX OF MULTIDISCIPLINARY NEURAL NETWORK SUPPORT OF MEDICAL DECISION MAKING IN THE TREATMENT OF CORONARY HEART DISEASE

DOI: 10.25881/18110193_2023_3_58

Abstract. *The article covers methods and procedures for developing a neural network decision support system when choosing the tactics of surgical intervention on coronary heart vessels. The system is designed to advise a wide range of practicing cardiologists and cardiac surgeons when deciding on the tactics of surgical intervention in patients with conditions associated with compromised coronary vessels. Based on a mathematical model taking into account a number of factors and the outcomes of previously performed surgeries, the neural network system offers a choice between aorto-coronary bypass surgery and percutaneous coronary intervention. The decision determined by the system can serve as an additional argument for the final adoption of a collegial decision in complex clinical cases. Right decision affects the patient's recovery time after surgery, the quality of life after recovery, and the ability to continue working after treatment. The neural network decision support system in the field of cardiac surgery is designed as a standard application for a personal computer with specific technical characteristics that allow processing a large amount of data. Access to the system can be obtained by any cardiologist or cardiac surgeon registered in the system and validated. The developed complex is designed to provide healthcare institutions with a digital product and domestic service based on a new technological structure.*

The work was supported by the Foundation for Assistance to Small Innovative Enterprises (FASIE).

Keywords: *cardiovascular diseases, diagnostics and treatment, ischemic heart disease, mathematical model, neural network, software.*

For citation: *Zhuravlev D.M., Kopylov F.Yu., Chaadaev V.K., Ardatov S.V., Chaadaev K.V. Automated complex of multidisciplinary neural network support of medical decisions in the treatment of coronary heart disease. Medical doctor and information technology. 2023; 3: 58-71. doi: 10.25881/18110193_2023_3_58.*

ВВЕДЕНИЕ

Согласно действующим рекомендациям, принятие решения как о необходимости реваскуляризации миокарда, так и о выборе метода (чрескожное коронарное вмешательство (ЧКВ) или аортокоронарное шунтирование (АКШ)) должно проводиться командой кардиологов и кардиохирургов с учетом оценки риска и пользы проводимой процедуры. При этом необходимо прогнозировать не только риск смерти в краткосрочном (внутригоспитальном) и долгосрочном периодах, но и оценивать качество жизни, ожидаемый функциональный класс заболевания, риск повторных инфарктов миокарда, вероятность повторных хирургических вмешательств.

Для стратификации риска разработаны различные шкалы:

- шкала EuroScore II учитывает 18 клинических факторов, в том числе возраст, пол, функцию почек, наличие сопутствующих заболеваний, перенесенных ранее хирургических вмешательств, функциональный класс сердечной недостаточности и стенокардии, сократимость миокарда, наличие легочной гипертензии, сроки проведения хирургического вмешательства, необходимость в дополнительном вмешательстве на аорте или клапанах сердца. Данная шкала прогнозирует риск только внутригоспитальной летальности при изолированном АКШ, а также сочетании АКШ с хирургическими вмешательствами на аорте или клапанах сердца. Таким образом, отсутствует оценка пользы от вмешательства в долгосрочном периоде, также шкала не подходит для расчета риска ЧКВ. Учитываются только клинические факторы, не оценивается анатомическое строение коронарного русла [1, 2];
- шкала Syntax Score применяется для выбора оптимального метода хирургического вмешательства (ЧКВ или АКШ) с учетом анатомического строения коронарного русла, при стратификации риска учитывается 11 факторов. Данная шкала позволяет прогнозировать риск/пользу ЧКВ в отсроченном периоде, но не АКШ. Не учитываются клинические факторы. Кроме того, шкала не применима для использования врачами кардиологами, расчет основан на визуальной оценке коронарного русла специалистом хирургом,

а следовательно, существует широкая вариация результатов с учетом субъективной оценки и опыта специалиста [3, 4];

- шкала Logistic Clinical SYNTAX совмещает в себя оценку как клинических, так и анатомических факторов. Однако при расчете риска учитывается только три клинических фактора: возраст, фракция выброса левого желудочка, креатинин. SYNTAX II учитывает 6 клинических и 12 анатомических факторов, позволяет оценить долгосрочные риски ЧКВ [5, 6].

Таким образом, существует ряд прогностических шкал, позволяющих стратифицировать риск у пациентов кардиохирургического профиля. Однако каждая из рассмотренных методик имеет ряд ограничений и недостатков: EuroSCORE II — универсальность применения у всех кардиохирургических больных, отсутствие показателей поражения коронарного русла; Syntax Score — объемность, отсутствие признаков поражения коронарного русла; SYNTAX II — ограниченное рассмотрение клинических показателей. Кроме того, эти шкалы не предназначены для проведения комплексной оценки как клинических факторов, так и анатомического строения коронарного русла, необходимой для выбора стратегии и тактики ведения пациентов с ишемической болезнью сердца. То есть, можно сказать, что на сегодняшний день существуют определенные проблемы с оценкой не только риска повторного инфаркта миокарда, но и качества жизни, а также функционального класса пациента после хирургического вмешательства.

В связи с этим возникает необходимость разработки интеллектуальной системы поддержки принятия решения при выборе тактики оперативного вмешательства на коронарных сосудах сердца на основе машинного обучения, предназначенной для консультирования широкого круга практикующих кардиологов и кардиохирургов.

Реваскуляризация миокарда является подходящим методом, если при принятии решения ожидаемая польза с точки зрения выживаемости и улучшения здоровья (симптомов, функционального статуса и/или качества жизни) превосходит ожидаемые негативные последствия процедуры. Решение о предпочтении

медикаментозного лечения, как и ЧКВ, АКШ, должно основываться на отношении риска польза соответствующих методов, включая риск смерти во время процедуры, инфаркта миокарда и инсульта в сопоставлении с улучшением качества жизни, продления жизни, удлинения периода без инфаркта или повторной реваскуляризации. Команда специалистов должна принимать во внимание анатомию коронарного русла, заболевание, возраст, сопутствующие заболевания, а также предпочтения пациента, навыки проводящего процедуру врача и общий опыт лечебного учреждения [7, 8].

Однако такой подход все же остается достаточно субъективным, значительно зависит от квалификации и накопленного практического опыта лечащего врача. Альтернативным вариантом выбора оптимального лечения в сложных клинических ситуациях являются информационные системы поддержки принятия врачебных решений с использованием нейронных сетей. Технологии искусственного интеллекта (англ. Artificial Intelligence, AI) способны существенно дополнить возможности человека в части получения более ясной картины рассматриваемой проблемы для повышения точности диагноза с использованием различных алгоритмов и когнитивных вычислений [9]. Основой технологии является машинное обучение (англ. Machine learning, ML), использующее модели, основанные на обучающих данных, необходимых для принятия решений, и программные алгоритмы для решения проблемы [10].

Авторами, после проведения анализа ряда математических методов [11–15], используемых для построения алгоритмов нейронных сетей, выбор был остановлен на комбинированной модели, включающей в себя:

- алгоритмы вычисления оценок, ключевым понятием которых является параметризованная функция близости;
- реализации генетического алгоритма для формирования векторов весов признаков (и групп признаков, например, весов объектов и медицинских характеристик) в целях классификации (разграничения опорных множеств) медицинских случаев (объектов), подбора весовых параметров функции близости и расширения обучающей выборки.

МЕТОДЫ И МАТЕРИАЛЫ

С прикладной точки зрения система поддержки принятия решения при выборе тактики оперативного вмешательства на коронарных сосудах сердца представляет собой информационную систему, состоящую из двух компонентов:

1. Подсистема обучения и управления знаниями, задача которой — обучение нейронной сети на массиве данных об оперированных пациентах. В формализованном описании каждого медицинского случая присутствуют такие сведения как: анамнез пациента, состояние здоровья до операции, протокол выполнения операции, информация о восстановлении пациента и наблюдение за его состоянием после операции.
2. Подсистема поддержки принятия решения, формирующая предложения по тактике лечения на основании знаний, заложенных в семантическую сеть. Чем более полно и точно будут описаны случаи, заложенные в базу знаний, тем более точное и квалифицированное решение может быть предложено системой.

В целом, при разработке информационной системы была применена следующая последовательность действий:

1. Формирование базы знаний:
 - разработка формальной конструкции описания случая лечения пациента (включая классификаторы, форматы ввода-вывода данных);
 - сбор и структурирование данных (из электронных медицинских карт, из текстовых документов, из экспертных анкет и пр.).
2. Опытная апробация системы поддержки принятия решений:
 - разработка формата запроса к системе, требований к интерфейсу и способам получения информации, то есть составление частного технического задания;
 - разработка механизма запросов к семантической сети, формированию решения и самообучения;
 - испытание системы в клинической практике.

В рамках решения задачи был сформирован массив из 26104 видеофайлов, содержащих результаты коронарографии более 1500 пациентов. Сбор данных по медицинским случаям, в том числе последующий обзвон прооперированных пациентов

выполнен специалистами Первого МГМУ им. И.М. Сеченова. Затем была проведена обработка данных на предмет их полноты, достоверности и степени структурированности сведений, характеризующих состояние пациента в динамике. Анализ проводился с применением апробированных технологий как в медицинской сфере [16, 17], так в других областях науки и техники [18]. По итогам расшифровки и обработки медиафайлов и данных были созданы 1000 карточек реальных медицинских случаев и 1 тестовая. 500 карточек были отсеяны по причине отсутствия в них сведений, необходимых для формирования полнотекстовой базы данных разработанного комплекса.

Результат заполнения карточки пациентов с опубликованными медиафайлами коронарографии приведен на рис. 1, 2 (приведена карточка условного пациента, используемая для отладочных и тестовых работ).

С точки зрения исследования медицинских случаев и подходов при принятии врачебных решений были формализованы факторы, риски и противопоказания при реваскуляризации для формирования комментариев информационной системы при выборе тактики лечения:

Блок Коронароангиографии:

- Абсолютных противопоказаний нет.
- Ограничения проведения коронарографии:
 - тяжелая почечная недостаточность: скорость клубочковой фильтрации ниже 30 (можно при наличии аппарата гемодиализа);
 - вес пациента более 130 кг (нежелательно более 120 кг);
 - известная анафилактическая реакция на йод;
 - нежелание пациента;
 - рецидивирующая желудочковая тахикардия;
 - выраженная тахисистолия при фибрилляции предсердий с нестабильным уровнем артериального давления.

Блок стентирования коронарных артерий:

- Абсолютных противопоказаний нет.
- Ограничения проведения коронарографии:
 - тяжелая почечная недостаточность: скорость клубочковой фильтрации ниже 30 (можно при наличии аппарата гемодиализа);

- вес пациента более 130 кг (нежелательно более 120 кг);
- известная анафилактическая реакция на йод;
- нежелание пациента;
- рецидивирующая желудочковая тахикардия;
- выраженная тахисистолия при фибрилляции предсердий с нестабильным уровнем артериального давления;
- диаметр артерии менее 2 мм.

Блок аортокоронарного шунтирования:

- Абсолютных противопоказаний к плановой АКШ нет.
- Оценка риска операции проводится по шкале EUROSCORE и EUROSCORE-2.
- Факторами риска периоперационной смертности являются:
 - женский пол;
 - хроническая болезнь почек II и более стадии;
 - патология периферических артерий;
 - перенесенные операции на сердце;
 - дыхательная недостаточность II и III степени;
 - острый эндокардит;
 - сахарный диабет;
 - хроническая сердечная недостаточность III и IV функционального класса;
 - фракция выброса левого желудочка ниже 50%;
 - перенесенный инфаркт миокарда;
 - легочная гипертензия умеренной и тяжелой степени.
- Относительными противопоказаниями к аортокоронарному шунтированию считаются:
 - диффузное поражение всех коронарных артерий;
 - снижение фракции выброса левого желудочка до 30%;
 - клинические признаки застойной сердечной недостаточности IV функционального класса;
 - дыхательная недостаточность тяжелой степени;
 - почечная или печеночная недостаточность тяжелого течения;
 - онкологические заболевания;
 - предполагаемая продолжительность жизни менее 3 лет.

Администрирование		Поддержка врачебных решений	
Алгоритм	Пациенты	Опросник	Поддержка решений
		Консилиумы	
00001 Редактировать Вернуться			
ФИО	Пробная Инесса Ивановна		
Пол	Женский		
Дата рождения	20.07.1957г. 60 лет		
Место жительства	Москва		
Телефон 1	+7(916)674-03-69		
Телефон 2			
Дата госпитализации	29.07.2017г.		
Вес	64 кг		
Рост	158 см		
ИМТ	25.64		
Стенокардия	Нет		
Острый инфаркт миокарда	Нет		
ПИКС	Повторные		
Вмешательства на коронарных сосудах	Нет		
Сахарный диабет	Второй тип		
Артериальная гипертензия	Скорректирована		
Фибрилляция предсердий	Нет		
Трепетание предсердий	Нет		
Желудочковая тахикардия	Нет		
Паузы более 3 секунд	Нет		
Курение	Нет		
Бронхиальная астма	Нет		
ХОБЛ	Нет		
ХСН	Нет		
Язва желудка	Нет		
Язва 12-перстной кишки	Нет		
Фракция выброса	60.0		
Толщина ЗСЛЖ	0.0 мм		
Толщина МЖП	0.0 мм		
Размер ПЖ	3.0 см		
Размер ЛП	5.0 см		
Объем ЛП	мл		
Конечный систолический объем	100.0 мл		
Конечный диастолический объем	50.0 мл		
Кальциоз аорты	Нет		
Аортальная регургитация	1		
Аортальный стеноз	1		
Митральная регургитация	1		
Митральный стеноз	1		
Систолическое давление в лёгочной артерии	мм рт.ст.		
Зоны акинеза	Нет		
Зоны гипокинеза	Нет		
Аневризма ЛЖ	Нет		
Аневризма грудного отдела аорты	Нет		
Заключение ЭхоКГ	дегенеративных признаков не выявлено		
Дата КАГ			
Тип кровоснабжения сердца	Сбалансированный		
Кол-во стенозов ствола ЛКА	2		
Стеноз ствола ЛКА	0.01%		
2-й стеноз ствола ЛКА	0.02%		
Окклюзия ствола ЛКА	Нет		
Кол-во стенозов проксимального сегмента ПМЖА	0		
ПМЖА стенозирована в устье	Нет		
Окклюзия проксимального сегмента ПМЖА	Нет		
Кол-во стенозов среднего сегмента ПМЖА	0		
Окклюзия среднего сегмента ПМЖА	Нет		
Кол-во стенозов дистального сегмента ПМЖА	0		
Окклюзия дистального сегмента ПМЖА	Нет		
Количество стенозов ветвей ПМЖА более 75%			
Окклюзии в ветвях ПМЖА	Нет		
Кол-во стенозов проксимального сегмента ИМА	3		
Стеноз проксимального сегмента ИМА	0.01%		
2-й стеноз проксимального сегмента ИМА	0.02%		
3-й стеноз проксимального сегмента ИМА	0.03%		
ИМА стенозирована в устье	Да		
Окклюзия проксимального сегмента ИМА	Да		
Кол-во стенозов среднего сегмента ИМА	3		
Стеноз среднего сегмента ИМА	0.21%		
2-й стеноз среднего сегмента ИМА	0.22%		
3-й стеноз среднего сегмента ИМА	0.23%		
Окклюзия среднего сегмента ИМА	Да		
Кол-во стенозов дистального сегмента ИМА	3		
Стеноз дистального сегмента ИМА	0.31%		
2-й стеноз дистального сегмента ИМА	0.32%		
3-й стеноз дистального сегмента ИМА	0.33%		
Окклюзия дистального сегмента ИМА	Да		
Кол-во стенозов проксимального сегмента ОА	0		
Окклюзия проксимального сегмента ОА	Нет		
Кол-во стенозов среднего сегмента ОА	0		

Рисунок 1 — Карточка пациента. Часть 1.

Стеноз брахиоцефальных сосудов	80.0	Окклюзия среднего сегмента ОА	Нет
Стеноз сосудов нижних конечностей	0.0	Кол-во стенозов среднего сегмента ОА	0
Заключение по стенозу сосудов МАГ		Окклюзия дистального сегмента ОА	Нет
Заключение по стенозу сосудов нижних конечностей		Количество стенозов ветвей ОА более 75%	4
Дата нагрузочного тестирования		Окклюзии в ветвях ОА	Нет
Нагрузочный тест - признаки ишемии	Нет	Кол-во стенозов проксимального сегмента ПКА	0
Антиангинальная терапия	Нет	Окклюзия проксимального сегмента ПКА	Нет
Максимальная ЧСС	в мин	Кол-во стенозов проксимального сегмента ПКА	0
Ишемия при ЧСС	в мин	Окклюзия среднего сегмента ПКА	Нет
Максимальное АД	мм рт.ст.	Кол-во стенозов проксимального сегмента ПКА	3
Толерантность к нагрузке	Низкая	Стеноз дистального сегмента ПКА	0.25%
Ишемия при нагрузке (METS)		2-й стеноз дистального сегмента ПКА	0.01%
Индукция желудочковой групповой аритмии	Нет	3-й стеноз дистального сегмента ПКА	0.02%
Заключение нагрузочной пробы	Пробная проба	Окклюзия дистального сегмента ПКА	Нет
Креатинин	1.11 мг/дл	Количество стенозов ветвей ПКА более 75%	7
СКФ	53.9324	Окклюзии в ветвях ПКА	Нет
Гемоглобин	120.0 г/л	Сохранность дистального коронарного русла	Умеренно изменено
АСТ	50.0 Ед/л	Заключение коронарографии	Тип кровоснабжения сердца - правый Ствол левой коронарной артерии: без R-признаков гемодинамически значимых стенозов Передняя межжелудочковая ветвь (ПМЖВ): окклюзия в проксимальном сегменте, дистальные отделы контрастируются фрагментарно по интракоронарным и межкоронарным коллатералям Огибающая артерия (ОА) без R-признаков гемодинамически значимых стенозов Правая коронарная артерия (ПКА): без K-признаков гемодинамически значимых стенозов.
АЛТ	50.0 Ед/л	Жив на момент выписки	Нет
Глюкоза	13.0 ммоль/л	Какую операцию выполнили	Ничего
НЬА1С	5.02%	Удалось дозвониться	Да, дата контакта:
Тромбоциты	121.0	Хирургическое вмешательство после КАГ	
Общий холестерин	10.0 ммоль/л	Инфаркт миокарда, связанный с ЧКВ/АКШ	
ЛПНП	5.0 ммоль/л	Инфаркт миокарда в течение последующих лет, не связанный с ЧКВ/АКШ	
		Повторные хирургические вмешательства на коронарных артериях	АКШ и ЧКВ Даты АКШ: 01.11.2017, 03.11.2017 Даты ЧКВ: 09.11.2017, 30.11.2017

Рисунок 2 — Карточка пациента. Часть 2.

При разработке математической модели поддержки врачебных решений при лечении ишемической болезни сердца были использованы следующие данные:

- сведенный и внесенный в базу массив данных по медицинским случаям, при этом для обучения сети было использовано 97%

выбранных клинических случаев и 3% клинических случаев для проверки адекватности обучения;

- алгоритмы вычисления оценок, обеспечивающих обучение на определенной изначальной выборке данных (формализованных в виде дискретных векторов значений) в

части принадлежности каждого вектора к одному из классов, позволяющей по «близости» вновь поступающего на обработку медицинского случая (вектора диагностических характеристик) к имеющимся в обучении векторам определить принадлежность медицинского случая к установленным классам;

- требования конечных потребителей информации, кардиохирургов, к предоставлению результатов работы.

Основными базовыми объектами математической модели стали:

- формализованная конструкция, описывающая случай обращения пациента, содержащий информацию о нем, на основании которой принимается решение (является объектом, записанным в виде характеристического вектора значений в терминах алгоритма вычисления оценок);
- классов объектов, определенный практикующими кардиохирургами, состоит из следующих элементов:
- пациент жив через 5 лет после проведения ЧКВ;
- пациент жив через 5 лет после проведения АКШ;
- смерть пациента через 5 лет после проведения ЧКВ;
- смерть пациента через 5 лет после проведения АКШ;
- невозможность рассчитать результат по АКШ;
- невозможность рассчитать результат по ЧКВ.

На данном этапе, фактически тестовой работы информационной системы, в качестве конечной точки взят временной интервал в 5 лет. Это период времени, в течение которого есть информация о состоянии пациентов в сформированных медицинских карточках. На основании данной информации система может предлагать ту или иную тактику лечения. При необходимости, в случае дальнейшего развития программного комплекса, временной интервал может быть увеличен, также может быть расширен и класс объектов;

- формализованная конструкция, описывающая результат проведения операции — успех или неуспех в перспективе как минимум в 5 лет, сформированный по результатам звона пациентов. Все результаты проведения

операции формируют опорные множества на фиксированных указанных выше классах;

- ряд модельных математических задач, а именно задачи классификации, теории алгоритмов вычисления оценок, генетических алгоритмов и функции близости (база вычислительного ядра; описывающее разработанное математическое и программное обеспечение).

РЕЗУЛЬТАТЫ

В рамках задачи по разработке программного приложения, реализующего приведенную выше математическую модель поддержки врачебных решений при лечении ишемической болезни сердца подсистемы поддержки принятия решения и подсистемы лечения пациента (подсистема обучения и управления знаниями), выполнены следующие действия:

- произведено развитие функциональной архитектуры и информационного обеспечения приложения;
- разработаны графические пользовательские интерфейсы подсистемы поддержки принятия решения и подсистемы обучения и управления знаниями.

Полнофункциональный прототип клиентской части приложения размещен в сети Интернет по адресу: <http://chd.molnet.ru/chd/cp> (логин и пароль — по запросу), на рис. 3 приведен пример пользовательского интерфейса (стартовая страница).

Для обучения разработанной нейросети был использован массив данных из 976 реальных медицинских случаев, по которым собраны порядка 150 характеристик (см. рис. 1–3). Обучение проводилось по следующему алгоритму:

- вычисление порогов решающих правил;
- построение функции близости, при помощи которой вычислялись веса признаков и объектов обучения;
- оптимизация системы опорных множеств;
- генерация векторов системы опорных множеств;
- оператор мутации применяется ко всем векторам системы средствами хранимой процедуры в базе данных, значение параметра оператора мутации берется равным 0,001 (эмпирическое значение применения генетического алгоритма для смежных задач);

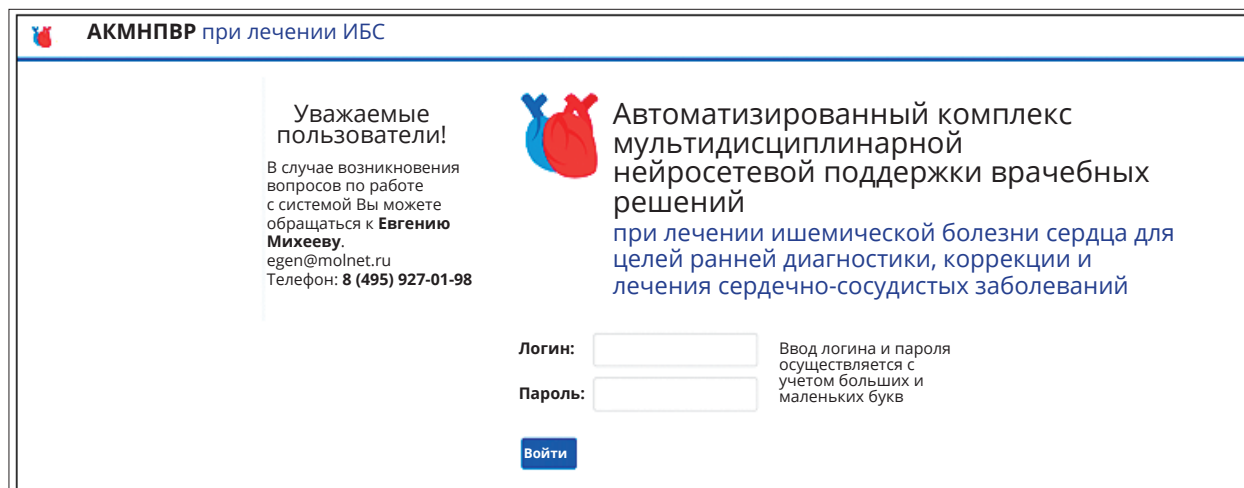


Рисунок 3 — Стартовая страница приложения.

- оценка каждого опорного множества за счет построения функции близости всех объектов контроля для каждого вектора из системы опорных множеств и занесение результатов в базу данных;
- генерация нового поколения с учетом оценки каждого вектора и занесение всех векторов системы в базу данных;
- после оценивания всей системы опорных множеств $\Omega\{A\}$ происходит проверка на критерий останова (выбрано 100 итераций работы алгоритма);
- проверка адекватности обучения.
Кроме того, в процессе обучения (или «цикл «скрещивание-мутация/селекция-размножение»» в терминах генетических алгоритмов) использовались итерационные технологии «обучения с учителем», в ходе которого система принудительно обучалась с помощью примеров «стимул-реакция». В итоге приложение произвело «самонастройку» параметров алгоритма вычисления оценок, также были выполнены мероприятия по проверке качества обучения.
Тестирование качества обучения нейросети было проведено на примерах, которые не участвовали в её обучении. Для чего из сформированного массива данных были взяты 24 реальных медицинских случая: 12 пациентов с выполненным ЧКВ и 12 — с выполненным АКШ.
Сравнение результатов работы приложения с тестовыми данными производилось по следующим категориям:
- по совпадению врачебных рекомендаций врача-хирурга и результата работы приложения:
 - совпало;
 - не совпало;
 - врач изменил собственное решение после ознакомления с результатом работы приложения («близкие» медицинские случаи и вероятности проведения операции);
- по результатам расчета врачебных рекомендаций приложением и исходам оперативного вмешательства:
 - врачебные рекомендации совпали с принятыми, пациент жив;
 - врачебные рекомендации совпали с принятыми, пациент умер;
 - врачебные рекомендации не совпали с принятыми, пациент жив;
 - врачебные рекомендации не совпали с принятыми, пациент умер.
 Указанные категории взаимно ортогональны, и каждая категория покрывает полностью все пространство вариантов, что позволяет построить матрицы сопоставления и затем сравнить их. Таким образом, была сформирована матрица результатов исследования, где по вертикали и горизонтали отложены указанные категории, а в ячейках — количество случаев, относящиеся к соответствующим категориям (см. табл. 1).
Ниже, на рисунках 4–7, визуализированы результаты расчета тактики лечения коронарных

Таблица 1 — Матрица результатов исследования

Врач, проводящий испытание, принял врачебное решение о тактике лечения пациента		Решение врача, проводящего испытание, в сравнении с решением, предложенным Системой:		
		совпало с решением, предложенным Системой	не совпало с решением, предложенным Системой	врач изменил собственное решение после ознакомления с решением, предложенным Системой
Решение врача, проводящего испытание, в сравнении с решением, принятым врачом, проводившим лечение пациента:	совпало с решением, принятым врачом, проводившим лечение пациента, и пациент жив	5	2	5
	совпало с решением, принятым врачом, проводившим лечение пациента, и пациент умер	1	1	-
	не совпало с решением, принятым врачом, проводившим лечение пациента, и пациент жив	3	3	2
	не совпало с решением, принятым врачом, проводившим лечение пациента, и пациент умер	-	-	2

Пояснения к таблице: красные ячейки — наиболее значимые ошибки алгоритма, требуемые уточнения. Зеленые ячейки — наиболее важные для оценки результативности работы Системы.

Язва 12-перстной кишки	Нет
Язва желудка	Нет
Положительный нагрузочный тест	Да
Отмена лечения	Да
Снижение сократимости миокарда при ЭХОКГ	Да
Артериальная гипертензия	Нет
Систолическое давление в лёгочной артерии	н/д
ХОБЛ	Нет
Бронхиальная астма	Нет
Стеноз сосудов нижних конечностей	н/д

Рисунок 4 — Данные пациента, часть 1.

сосудов сердца одного из пациентов контрольной группы.

Также необходимо отметить, что малая выборка базы исследования (1000 реальных случаев) не позволяет говорить о строгой

рекомендации использовать приложение, поэтому в ходе дальнейших натурных испытаний предполагается проведение полномасштабного обучения нейронной сети приложения.

Фибрилляция предсердий	Нет
Аортальная регургитация	0
Стенокардия	Стабильная III класса
ХСН	ФК по NYHA II
Ишемия при нагрузке (METs)	3.0
Аневризма ЛЖ	Нет
Толщина МЖП	н/д
Толщина ЭСЛЖ	н/д
ПИКС	Есть
ОИМ	Нет

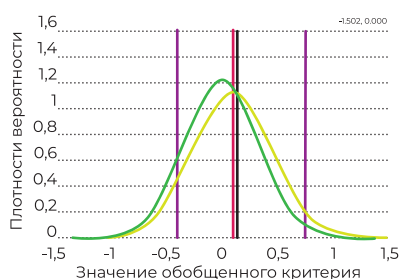
Рисунок 5 — Данные пациента, часть 2.

Митральная регургитация	0
Митральный стеноз	0
Аортальный стеноз	0
Фракция выброса	60.0
ЧСС при ишемии	100 в мин.
Сохранность дистального русла	Нет
Стенозы ветвей ОД	10.0%
Стенозы ветвей ПКА	10.0%
Стенозы ветвей ПМЖА	10.0%
Стенозы ИМА	10.0%

Рисунок 6 — Данные пациента, часть 3.

Результаты расчета вероятности наличия ИБС у пациента	Диагноз подтвержден 100%			
Результаты расчета тактики лечения коронарных сосудов сердца Значение обобщенного критерия (ОК): 0,139	<table border="1"> <tr> <td>Чрескоронарное вмешательство Надежность решения 49% Высокая</td> <td>Консервативное лечение</td> <td>Аортокоронарное шунтирование 0%</td> </tr> </table>	Чрескоронарное вмешательство Надежность решения 49% Высокая	Консервативное лечение	Аортокоронарное шунтирование 0%
Чрескоронарное вмешательство Надежность решения 49% Высокая	Консервативное лечение	Аортокоронарное шунтирование 0%		

Визуализация результатов



Условные обозначения

- Распределение консервативного лечения
- Распределение ЧКВ
- Распределение АКШ
- Значение обобщенного критерия для пациента
- Величины порогов

Краткое описание работы алгоритма

Результатом работы диагностического комплекса (ДК) должно стать решение двух задач:
 1. Обоснование достоверности реализованного в ДК подхода (методического и алгоритмического) путем сопоставления количества выбранных по определенному критерию тактик лечения (ТЛ) больных с реально имеющимися в базе данных (БД) ДК для обучающих и контрольных выборок;
 2. Предложения по улучшению выбора ТЛ больных путем изменения градаций критерия выбора тактики лечения (КВТЛ)

Рисунок 7 — Результаты расчета тактики лечения коронарных сосудов сердца.

ОБЛАСТЬ ПРИМЕНЕНИЯ РЕЗУЛЬТАТОВ

Основным функциональным назначением разработанного мультидисциплинарного комплекса информационной поддержки врачебных решений при лечении ишемической болезни сердца является консультирование широкого круга практикующих кардиологов и кардиохирургов при принятии решения о тактике проведения оперативного вмешательства при состояниях пациентов, связанных с нарушением проводимости коронарных сосудов.

ОБСУЖДЕНИЕ

Использование цифровых технологий в медицине, в частности искусственного интеллекта, позволяет создать доступные и недорогие в эксплуатации высокоточные диагностические инструменты для скрининга и выявления у пациентов различных хронических заболеваний. Более того, такие технологии могут значительно помочь сократить время и стоимость постановки диагноза, тем самым повысить пропускную способность учреждений здравоохранения [19].

Но, прежде всего, перспективы систем с использованием искусственного интеллекта заключаются в возможностях технологии обрабатывать большие объемы данных и фиксировать закономерности, которые традиционным способом было бы трудно идентифицировать. Это позволяет существенно снизить нагрузку на практикующих специалистов, минимизируя усталость и переутомление врачебного персонала. Именно поэтому за подобными технологиями стоит будущее, которое, при правильном подходе, способно кардинальным образом изменить практику медицины. Из приведенных в обзоре [20] примеров явно виден высокий интерес мирового медицинского сообщества к вопросу использования нейронных сетей для диагностирования патологических состояний. Методы системного анализа и создание систем интеллектуальной поддержки при соответствующей проработке могут быть применены при

описании практически любого класса заболеваний.

В настоящее время проводятся работы по расширению функциональности разработанного комплекса и оценка возможностей его использования в ортопедии на базе Самарского государственного медицинского университета.

ВЫВОДЫ

Полученные в ходе тестирования разработанного мультидисциплинарного комплекса информационной поддержки врачебных решений при лечении ишемической болезни сердца результаты дают основания, что цифровой продукт может быть востребован медицинским сообществом, что, в свою очередь, позволит сформировать запросы на дальнейшее развитие приложения по следующим направлениям:

- расширение результатов формирования врачебных назначений за счет обеспечения возможности рекомендаций вида «показания к медикаментозному лечению» и «врачебное вмешательство не рекомендовано, показаны физические методы лечения»;
- подготовка полного перечня, рассчитанного математическими методами, значимых характеристик при выборе метода лечения для последующего анализа и выявления закономерностей в целях расширения медицинского знания на более глубоком уровне;
- проведение значительного количества апробаций системы научным сообществом как для повышения достоверности рассчитанных решений, так и для повышения уверенности практикующих врачей в результирующих рекомендациях и перестроения технологического уклада при диагностике пациентов с ишемической болезнью сердца;
- расширение возможностей приложения за счет внедрения технологий искусственного интеллекта распознавания поврежденных сосудов при коронарографии.

ЛИТЕРАТУРА/REFERENCES

1. Gogayeva O. Comparison of the EuroSCORE I, EuroSCORE II and STS Scales in Cardiac Surgery of High-Risk Patients with Complicated Forms of CAD. Ukrainian journal of cardiovascular surgery. 2020; 3(40): 15-21.

2. Singh N., Gimpel D., Parkinson G. et al. Assessment of the EuroSCORE II in a New Zealand Tertiary Centre. *Heart, Lung and Circulation*. 2019; 28(11): 1670-1676. doi: 10.1016/j.hlc.2018.09.004.
3. Hatamnejad MR, Heydari AA, Salimi M, et al. The utility of SYNTAX score predictability by electrocardiogram parameters in patients with unstable angina. *BMC Cardiovasc Disord*. 2022; 22(8). doi: 10.1186/s12872-022-02455-6.
4. Minamisawa M, Miura T, Motoki H, et al. Prediction of 1-year clinical outcomes using the SYNTAX score in patients with prior heart failure undergoing percutaneous coronary intervention: sub-analysis of the SHINANO registry. *Heart Vessels*. 2017; 32(4): 399-407.
5. Modolo R, Collet C, Onuma Y, Serruys PW. SYNTAX II and SYNTAX III trials: what is the take home message for surgeons? *Annals of Cardiothoracic Surgery*. 2018; (4): 470-483. doi: 10.21037/acs.2018.07.02.
6. Nibber A, Ziada KM, Whayne TF. Predictive Value of the Logistic Clinical SYNTAX Score. *Angiology*. 2015; 66(8): 711-713. doi: 10.1177/0003319714562244.
7. Самородская И.В., Чернявская Т.К., Какорина Е.П., Семенов В.Ю. Ишемические болезни сердца: анализ медицинских свидетельств о смерти // Российский кардиологический журнал. — 2022. — Т.27. — №1. — С.22-28. [Samorodskaya IV, Chernyavskaya TK, Kakorina EP, Semenov VYu. Ishemicheskie bolezni serdca: analiz medicinskix svideteľ`stv o smerti. Rossijskij kardiologicheskij zhurnal. 2022; 27(1): 22-28. (In Russ.)]
8. Чернявская Т.К., Глезер М.Г. Клиническая характеристика и лечение амбулаторных пациентов с хронической сердечной недостаточностью в Московской области // Альманах клинической медицины. — 2021. — Т.49. — №2. — С.125-131. [Chernyavskaya TK, Glezer MG. Klinicheskaya kharakteristika i lecheniye ambulatornykh patsiyentov s khronicheskoy serdechnoy nedostatochnost'yu v Moskovskoy oblasti. Al'manakh klinicheskoy meditsiny. 2021; 49(2): 125-131. (In Russ.)]
9. Romiti S., Vinciguerra M., Saade W., Anso Cortajarena I., Greco E. Artificial Intelligence (AI) and Cardiovascular Diseases: an Unexpected Alliance. *Cardiology Research and Practice*. 2020; 2020: 1-8.
10. Zghyer F, Yadav S, Elshazly MB. Artificial Intelligence and Machine Learning. *Precision Medicine in Cardiovascular Disease Prevention*. 2021; 18: 133-148.
11. Созыкин А.В. Обзор методов обучения глубоких нейронных сетей // Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика. — 2017. — Т.6. — №3. — С.28-59. [Sozykin AV. Obzor metodov obucheniya glubokikh neyronnykh setey. Vestnik Yuzhno-Ural'skogo gosudarstvennogo universiteta. Seriya: Vychislitel'naya matematika i informatika. 2017; 6(3): 28-59. (In Russ.)]
12. Shaker M, Tantawi M, Shedeed A, Tolba F. Generalization of convolutional neural networks for ECG classification using generative adversarial networks. *IEEE Access*. 2020; 8: 35592-35605.

13. Голубев А.С., Звягин М.Ю., Прокошев В.Г., Рожков М.М. Особенности распознавания методом ближайшего элемента в алгоритмах вычисления оценок // Прикладная информатика. — 2013. — №1(43). — С.87-94. [Golubev AS, Zvyagin MYU, Prokoshev VG, Rozhkov MM. Osobennosti raspoznavaniya metodom blizhayshego elementa v algoritmakh vychisleniya otsenok. Prikladnaya informatika. 2013; 1(43): 87-94. (In Russ.)]
14. Журавлев Ю.И., Назаренко Г.И., Рязанов В.В., Клейменова Е.Б. Новый метод анализа риска развития ишемической болезни сердца с использованием геномных и компьютерных технологий // Кардиология. — 2011. — Т.51. — №2. — С.19-25. [Zhuravlev YUI, Nazarenko GI, Ryazanov VV, Kleymenova YeB. Novyy metod analiza riska razvitiya ishemicheskoy bolezni serdtsa s ispol'zovaniyem genomnykh i komp'yuternykh tekhnologiy. Kardiologiya. 2011; 51(2): 19-25. (In Russ.)]
15. Son C, Kim Y, Kim H, Park H, Kim M. Decision-making model for early diagnosis of congestive heart failure using rough set and decision tree approaches. Journal of Biomedical Informatics. 2012; 45: 999-1008.
16. Высоцкая Е.В., Беловод А.Н., Кириченко А.Н. Восстановление пропущенных значений параметров в таблицах биохимических анализов пациентов с псориазом // Вестник Национального технического университета Харьковский политехнический институт. Серия: Информатика и моделирование. — 2010. — №21. — С.38-45. [Vysotskaya EV, Belovod AN, Kirichenko AN. Renewal of the skipped values of parameters in the tables of biochemical analyses of patients with psoriasis. Herald of the National Technical University «KhPI». Subject issue: Information Science and Modeling. 2010; 21: 38-45. (In Russ.)]
17. Lipton ZC, Kale D, Wetzels R. Directly modeling missing data in sequences with rnns: Improved classification of clinical time series. In Machine Learning for Healthcare Conference. 2016; 253-270.
18. Шафроненко А.Ю., Волкова В.В., Бодянский Е.В. Адаптивная кластеризация данных с пропущенными значениями // Радиоэлектроника, информатика, управление. — 2011. — №2(25). — С.115-119. [Shafronenko AYU, Volkova VV, Bodyanskiy YeV. Adaptivnaya klasterizatsiya dannykh s propushchennymi znacheniyami. Radioelektronika, informatika, upravleniye. 2011; 2(25): 115-119. (In Russ.)]
19. Астахова И.Ф., Киселева Е.И. Интеллектуальная поддержка принятия врачебных решений // Современные информационные технологии и ИТ-образование. — 2020. — Т.16. — №3. — С.664-672. [Astakhova IF, Kiseleva EI. Intellektual'naya podderzhka prinyatiya vrachebny`x reshenij. Sovremennyy`e informacionny`e tekhnologii i IT-obrazovanie. 2020; 16(3): 664-672. (In Russ.)]
20. Sharma M, Savage C, Nair M, Larsson I, Svedberg P, Nygren JM. Artificial Intelligence Applications in Health Care Practice: Scoping Review. J Med Internet Res. 2022; 24(10): e40238.

КЛИМКО В.И.,

к.т.н., ФГБУ «ЦНИИОИЗ» Минздрава России, Москва, Россия,
e-mail: klimkovi@minzdrav.gov.ru

СОЛОМАТНИКОВ И.А.,

ФГБУ «ЦНИИОИЗ» Минздрава России, Москва, Россия,
e-mail: solomatnikovia@mednet.ru

ХОДАКОВА О.В.,

д.м.н., доцент, ФГБУ «ЦНИИОИЗ» Минздрава России, Москва, Россия,
e-mail: khodakovaov@mednet.ru

КУНГУРЦЕВ О.В.,

ФГБУ «ЦНИИОИЗ» Минздрава России, Москва, Россия,
e-mail: kungurtsevov@mednet.ru

СТРАДЫМОВ Ф.И.,

ФГБУ «ЦНИИОИЗ» Минздрава России, Москва, Россия,
e-mail: stradymovfi@mednet.ru

ЧЕРНОМЫРДИНА Е.В.,

ФГБУ «ЦНИИОИЗ» Минздрава России, Москва, Россия,
e-mail: chernomyrdinaev@mednet.ru

ИСПОЛЬЗОВАНИЕ КАРТОГРАФИЧЕСКИХ СХЕМ ТЕРРИТОРИАЛЬНОГО ПЛАНИРОВАНИЯ В ЗДРАВООХРАНЕНИИ

DOI: 10.25881/18110193_2023_3_72

Аннотация.

В статье рассматривается опыт применения инструментов визуализации данных при подготовке схем территориального планирования региональных систем здравоохранения для принятия управленческих решений по увеличению доступности медицинской помощи. Описан опыт формирования картографических данных для анализа текущего состояния и целевого уровня различных систем оказания медицинской помощи региона по профилям. Приведены тенденции развития средств визуализации, а также определена необходимость наличия специфических компетенций по работе с визуальными картографическими данными у лиц, принимающих решения (далее – ЛПР). Визуализация схем территориального планирования систем здравоохранения регионов позволяет системно оценить организацию оказания медицинской помощи на основе картографического анализа и информации о взаимном расположении медицинских организаций, оперативно получить объективные данные о доступности медицинской помощи населению и ускорить процесс принятия управленческих решений, направленных на устранение выявленных проблем. Всё это требует наличия у ЛПР сформированных компетенций визуальной коммуникации и визуальной грамотности.

Ключевые слова: визуализация данных, принятие управленческих решений, здравоохранение, BI, геоданные, business intelligence.

Для цитирования: Климко В.И., Соломатников И.А., Ходакова О.В., Кунгурцев О.В., Страдымов Ф.И., Черномырдина Е.В. Использование картографических схем территориального планирования в здравоохранении. Врач и информационные технологии. 2023; 3: 72-83. doi: 10.25881/18110193_2023_3_72.

KLIMKO V.I.,

PhD, Central Research Institute of Organization and Informatization of Healthcare of the Ministry of Health of Russia, Moscow, Russia, e-mail: klimkovi@minzdrav.gov.ru

SOLOMATNIKOV I.A.,

Central Research Institute of Organization and Informatization of Healthcare of the Ministry of Health of Russia, Moscow, Russia, e-mail: solomatnikovia@mednet.ru

KHODAKOVA O.V.,

DSc, Associate Professor, Central Research Institute of Organization and Informatization of Healthcare of the Ministry of Health of Russia, Moscow, Russia, e-mail: khodakovaov@mednet.ru

KUNGURTCEV O.V.,

Central Research Institute of Organization and Informatization of Healthcare of the Ministry of Health of Russia, Moscow, Russia, e-mail: kungurtsevov@mednet.ru

STRADYMOV F.I.,

Central Research Institute of Organization and Informatization of Healthcare of the Ministry of Health of Russia, Moscow, Russia, e-mail: stradymovfi@mednet.ru

CHERNOMYRDINA E.V.,

Central Research Institute of Organization and Informatization of Healthcare of the Ministry of Health of Russia, Moscow, Russia, e-mail: chernomyrdinaev@mednet.ru

USING CARTOGRAPHIC SCHEME OF TERRITORIAL PLANNING IN HEALTH CARE

DOI: 10.25881/18110193_2023_3_72

Abstract. *This article discusses the experience of using data visualization tools in the preparation of territorial planning schemes for regional health systems. The article also covers trends in the development of visualization tools and the competencies necessary for decision makers to use these tools effectively in making managerial decisions.*

Keywords: *data visualization, management decision-making, primary health care, BI-systems.*

For citation: *Klimko V.I., Solomatnikov I.A., Khodakova O.V., Kungurtcev O.V., Stradymov F.I., Chernomyrdina E.V. Using cartographic scheme of territorial planning in health care. Medical doctor and information technology. 2023; 3: 73-83. doi: 10.25881/18110193_2023_3_72.*

ВВЕДЕНИЕ

Технологическое развитие ускоряет накопление сведений об окружающем мире. Рост и накопление объема информации далеко не всегда приводит к принятию основанных на них решений. Согласно исследованию Bloomberg 80% руководителей корпораций не используют все данные, которые собирают [1]. Растущий объем информации может быть обработан с помощью технологий искусственного интеллекта. Опрос, проведенный в 2022 году ФГБУ «Центральный научно-исследовательский институт организации и информатизации здравоохранения» Министерства здравоохранения Российской Федерации (далее — ЦНИИОИЗ), показал, что руководители в сфере здравоохранения позитивно смотрят на внедрение искусственного интеллекта: 89% респондентов считают, что он полезен для медицины [2]. В этой цепочке для человека остается важная роль — принятие решений, основанных на анализе данных, полученных по итогам деятельности организации. Для этого лицам, принимающим решения, (далее — ЛПР) необходимы соответствующие инструменты визуализации, позволяющие доступным образом и в сжатый срок изучить имеющиеся данные и результаты их обработки для принятия эффективных решений.

В настоящее время происходят активные трансформационные преобразования в работе с медицинской статистикой и данными о ресурсных потребностях медицинских организаций (МО). Пандемия новой коронавирусной инфекции COVID-19 ускорила реализацию мероприятий по цифровизации предоставления статистических данных: введена новая модель информационного взаимодействия по удаленному согласованию и обработке данных форм годового отчета, внедрены новые формы ежемесячных отчетов и суточных мониторинговых форм, а также организован сбор данных Ad Hoc (метод создания отчетов из базы данных на основе требований и запросов пользователя), что приводит к росту их объема с последующей обработкой и анализом для принятия решений [3]. Средства визуализации позволяют провести скрининг гипотез и представить результаты проведенного анализа в удобном для восприятия ЛПР виде.

Современные BI-системы (Business Intelligence с английского «бизнес-аналитика») и облачные инструменты позволяют создавать визуальные решения Ad Hoc практически любого уровня сложности. Важность быстрой обработки и демонстрации полученных результатов для принятия и корректировки принятых решений представлена в публикациях по вопросам влияния геопространственной визуализации на принятие решений для борьбы с COVID-19, например, в Шеньчжэне [4], США [5] и России [6, 7].

Средства и возможности визуализации данных демонстрируют активное развитие. Согласно отчету Fortune Business Insights (2022) рынок ПО в сфере визуализации информации за период 2023–2030 гг. будет расти в среднем на 9,1% в год [8]. Развитие искусственного интеллекта существенно мультиплицирует возможности обработки данных, описания полученных результатов и в целом ускорение проверки гипотез. Важным трендом станет умение применять соответствующие средства сбора, обработки и визуализации данных при анализе ситуации. В процессы непосредственного оказания медицинской помощи уже начинают внедряться решения, основанные на искусственном интеллекте и при активном использовании средств визуализации [9, 10]. Такое же будущее может ждать и сферу организации здравоохранения [11].

При увеличении объема данных происходит интенсификация скорости принятия решений. Понятное представление результатов анализа данных позволяет повысить скорость принятия решений в вопросах здоровья населения на уровне ЛПР, что отмечают российские эксперты [12, 13], и эти тенденции будут сохраняться. Следовательно, необходимо адаптироваться под более высокие темпы цифровизации и развивать отечественные аналитические системы, включая системы визуализации данных и BI-аналитики.

Визуализация геоданных и сопутствующей информации для целей принятия решений в сфере здравоохранения осуществляется как на региональном уровне посредством применения программного обеспечения с открытым кодом или собственных разработок [14], так и на федеральном уровне путем использования

геоинформационных систем для типовых задач. Так, для территориального планирования в здравоохранении используется соответствующая геоинформационная система Минздрава России [15].

Растущие темпы принятия решений приводят к необходимости дальнейшего совершенствования инструментов визуализации данных и их адаптации под актуальные задачи. Также необходимо учитывать региональную специфику при принятии решений на уровне субъекта РФ. Таким образом, возрастает запрос на создание индивидуальных решений на региональном уровне.

Указ Президента РФ от 06.06.2019 №254 «О Стратегии развития здравоохранения в Российской Федерации на период до 2025 года» (далее — Стратегия) определил в числе основных задач развития здравоохранения в РФ создание условий для повышения доступности и качества медицинской помощи [16]. Проблемы доступности медицинской помощи возможно выявить при проведении картографического анализа существующей сети МО субъекта РФ на основе строгого выполнения положений всех нормативных правовых актов, регламентирующих территориальное планирование в здравоохранении [17].

Визуализация схем территориального планирования региональных систем здравоохранения позволяет разработать мероприятия по устранению возможных дефектов в организации медицинской помощи населению и сформировать целевые модели территориального планирования, обеспечивающие реализацию Стратегии.

При анализе и экспертной оценке схем территориального планирования ЛПР должны обладать определенными навыками и компетенциями по работе с визуальной информацией. При текущем уровне развития технологий и общества ответственность за принимаемые решения невозможно переложить целиком на автоматизированные системы с искусственным интеллектом, поскольку принятие решений осуществляется человеком, как объектом, наделенным свободой воли [18].

Для минимизации доли решений, продиктованных эмоциями и интуицией, ЛПР необходимо обладать соответствующими

компетенциями для корректного применения инструментов визуализации данных в рутинной практике. Для эффективного формирования позиции по возможным решениям ЛПР необходимо также обладать навыками визуальной коммуникации и визуальной грамотности. Под последними понимается способность воспринимать, интерпретировать и наделять смыслом информацию, представленную в визуализированной форме [19].

Отдельно стоит отметить навык выявления и интерпретации искажений, которые привносит визуализация данных. В отсутствие должных компетенций и опыта у ЛПР может формироваться искаженная картина, которая приведет к принятию нерациональных решений. Так, зарубежные авторы подчеркивают среди ключевых факторов недостаточную прозрачность процесса получения и обработки данных [20], отсутствие доказанной причинности и даже прямое манипулирование пользователями через инструменты визуального восприятия, используемых в политических целях [21].

Среди ключевых ошибок, способных привести к неоднозначному принятию решений при анализе карт с визуализацией данных, отметим типовые искажения:

1. *Однозначная трактовка цветовой палитры при анализе геоданных на плоскости в реальном масштабе.* При формировании «тепловых» карт или иной инфографики, относящейся к геоданным, лучше избегать цветowych решений «светофорного» типа: красный — плохо, зеленый — хорошо. В таких случаях неизбежны искажения масштаба той или иной ситуации: при наличии проблемы в муниципальном районе, занимающем значительную долю по площади территории, ситуация будет казаться негативной для всего региона в целом, хотя площадь района зачастую является второстепенным показателем после численности населения.
2. *Изменение масштаба при формировании карт в проекции Меркатора.* Эффект проявляется при визуализации данных на уровне страны в целом. Отображение земной поверхности на плоскости неизбежно искажает реальные размеры территорий, и

истинная площадь региона не может быть напрямую сопоставима.

3. *Сложности сравнения при использовании неконтрастных цветовых решений.* При выборе цветовой палитры итоговая контрастность оцениваемого и соседнего объектов (полигонов) может быть недостаточной, что приведет к принижению значимости объекта при разработке решения. Аналогичным образом будет хуже восприниматься информация без четких контурных линий: очертания полигонов могут сливаться как с фоном, так и друг с другом.

Корректное использование средств визуализации и развитие навыков анализа визуальной информации у ЛПР позволит распознать искажения, обеспечит корректную трактовку аналитического результата и снизит риск принятия нерациональных решений.

МАТЕРИАЛЫ И МЕТОДЫ

С января 2023 года ЦНИИОИЗ совместно с региональными органами исполнительной власти в сфере охраны здоровья проводит работу по созданию региональных схем территориального планирования в здравоохранении. Авторами были проанализированы схемы территориального планирования Пермского края, Тамбовской области, Республики Ингушетия, разработанные в рамках реализации действующих нормативно-правовых актов [22–25].

Для визуализации транспортной доступности объектов здравоохранения применялась карта региона, на которую были нанесены ключевые элементы, необходимые для принятия управленческих решений.

Географические карты с указанием границ районов для работ использовались в формате GeoJSON. Это стандартизированный формат представления географических структур данных, основанный на JSON (англ. JavaScript Object Notation), подходящий для целей визуализации геопространственных данных на уровне выделения полигонов — муниципальных округов, регионов.

В качестве входных данных для создания картографических материалов использовались excel-таблицы с адресами учреждений по профилям оказания медицинской помощи в

субъектах и .shp-файлы с полигональной геометрией границ субъектов, выгружаемые из сервиса OpenStreetMap при помощи плагина для ГИС QGIS «QuickOSM». Полученные данные загружались в проект ГИС ArcGIS Pro, создавались точечные слои с МО по профилям, найденными по адресам с помощью поиска по встроенным в ГИС локаторам. Далее производился расчет времени доезда от административного центра каждого муниципального образования по данным сервиса Яндекс.Карты с поправкой на наличие пробок: если на маршруте есть железнодорожные переезды, загруженные перекрестки или другие значимые препятствия, в качестве времени использовалось среднее между наибольшим и наименьшим возможным временем доезда. После проведенных вычислений в атрибутивную таблицу .shp-файла добавлялись поля (столбцы) для каждого профиля, и ячейки таблицы заполнялись имеющимися данными. Затем через меню «Символы» полигонам присваивались цвета, соответствующие времени доезда по профилю оказания медицинской помощи, и создавалась компоновка карты — добавлялись название, легенда, дополнительные текстовые сведения. Расчеты и оформление карт выполнялись сперва по всем исследуемым профилям для текущей ситуации, затем для планового состояния на 2025 год.

Для визуализации объектов здравоохранения региона и выработки управленческих решений по повышению доступности медицинской помощи, разрабатывались картографические изображения отдельных систем оказания медицинской помощи по профилям, под которыми понимаются объединенные единством функций или подходов в работе с пациентами по профилям МО, их подразделения, клинические и параклинические службы, участвующие в оказании медицинской помощи населению.

При оценке доступности медицинской помощи для отдельных элементов систем здравоохранения были сформированы индивидуальные карты:

- система оказания помощи пациентам с сосудистыми катастрофами;
- система оказания помощи пациентам с онкологическими заболеваниями;

- система оказания помощи беременным, роженицам и новорожденным;
- система оказания помощи пострадавшим в дорожно-транспортном происшествии (ДТП);
- система оказания медицинской помощи детскому населению;
- система оказания первичной медико-санитарной помощи (ПМСП);
- система оказания скорой медицинской помощи (СМП).

Для упрощения выработки управленческих решений картографический анализ был дополнен целевой моделью оказания медицинской помощи по профилям и условиям оказания медицинской помощи.

Целевая модель по условиям и профилям оказания медицинской помощи является главным инструментом, направленным на характеристику системы регионального здравоохранения по двум основным параметрам: коечному фонду и территориальному расположению МО и их структурных подразделений.

Поскольку объемы коечного фонда определяются с учетом множества корректирующих параметров, а объемы оказываемой медицинской помощи в стационарных условиях формируются с учетом региональных поправочных коэффициентов, то полученные отклонения не могут лежать в основе реорганизационных мероприятий, но позволяют определить вектор анализа и совершенствования системы здравоохранения субъекта.

При определении профицита фактического количества коек к расчетному рассматривалась маршрутизация пациентов по профилям, эффективность использования имеющихся коек, уровень госпитализации и больничная летальность. При определении дефицита фактического количества коек к расчетному, вырабатывались мероприятия по созданию дополнительных стационарных мощностей. Отклонение фактического и расчетного числа коек оценивалось по двум основным критериям: эффективность работы коечного фонда и расположение коек по профилю, исходя из требований к доступности.

Мероприятия по достижению целевой модели для рассматриваемых регионов разработаны исходя из наименьших затрат, включая

временные затраты. Мероприятия схем территориального планирования являются основой для формирования региональных программ и проектов, поэтому они не содержат сроков исполнения, характеристик критериев достижения результатов и ответственных за реализацию. Оцениваемые схемы территориального планирования и включенные в нее мероприятия являются следствием комплексной оценки множества факторов и позволяют организовать стратегическое управление здравоохранением региона на основе данных.

РЕЗУЛЬТАТЫ

При анализе содержания типового объекта визуализации были выявлены следующие паттерны и определены возможности для оптимизации.

На карте региона были выделены и подписаны муниципальные районы. При работе с картой в распечатанном виде или в виде выгрузки из VI-системы наименование района отображалось непосредственно на карте. Оптимальным способом являлось нанесение наименования внутри границ района. Если графический элемент не позволял уместить наименование, допускались сокращения или выноски наименования.

Каждый из районов должен иметь наглядную дифференциацию для оценки базового состояния с учетом специфических для каждой системы показателей. Так, например, для системы оказания помощи пациентам с сосудистыми катастрофами использовались следующие показатели:

- численность взрослого населения, тыс. чел.;
- плотность взрослого населения, чел. на км²;
- смертность населения от болезней системы кровообращения, на 100 тыс. населения;
- время доезда СМП, мин.

Для цветовой дифференциации применялись: значения медианы, верхний предел первого квартиля (Q1) и верхний предел третьего квартиля (Q3). Таким образом, была получена четырехцветная шкала по пороговым значениям квартиля: менее Q1 — цвет №1, от Q1 до медианы — цвет №2, от медианы до Q3 — цвет №3, от Q3 и выше — цвет №4. Для снижения искажений, связанных с цветовым восприятием карты, отдельные значения показателей

наносились непосредственно на карту. При нанесении числовые значения показателей округлялись до двух значимых цифр.

Для разных карт использовалась одинаковая цветовая палитра согласно логике формирования цветовой шкалы. Это позволило

избежать цветового зашумления. Тем не менее, при составлении нескольких карт с оценкой доступности в различных аналитических разрезах, с целью обеспечить четкую дифференциацию рассматриваемых изображений между собой, допускались исключения в виде

Таблица 1 — Основные элементы системы по профилю медицинской помощи

№ п/п	Система	Элементы
1	Система оказания помощи пациентам с сосудистыми катастрофами	РСЦ для ОНМК (3 уровень) РСЦ для ОКС (3 уровень) ПСО для ОНМК (2 уровень) ПСО для ОКС (2 уровень) неврологические койки (1 уровень) кардиологические койки (1 уровень) Наличие КТ в МО
2	Система оказания помощи пациентам с онкологическими заболеваниями	Онкологический диспансер (3 уровень) ЦАОП (2 уровень) Первичные онкологические кабинеты (1 уровень)
3	Система оказания помощи беременным, роженицам и новорожденным	Перинатальный центр (3 уровень) Родильное отделение (3 уровень) Родильное отделение (2 уровень) Родильное отделение (1 уровень) Женская консультация (1 уровень) Кабинет акушера-гинеколога
4	Система оказания помощи пострадавшим в ДТП	Многопрофильная МО (1 уровень) ТСО (2 уровень) ХСО (3 уровень)
5	Система оказания медицинской помощи детскому населению	Детская больница Детское отделение
6	Система оказания ПМСП	Отделение /кабинет ВОП ФАП/ФП Врачебная амбулатория Поликлиническое подразделение
7	Система оказания СМП	Пункт базирования санавиации Станция СМП Отделение СМП Подстанция СМП Число автомобилей СМП по категориям (А, В, С)

Сокращения: * РСЦ — региональный сосудистый центр; ОНМК — острое нарушение мозгового кровообращения; ПСО — первичное сосудистое отделение; ОКС — острый коронарный синдром; КТ — компьютерный томограф; ДТП — дорожно-транспортное происшествие; МО — медицинская организация; ЦАОП — центр амбулаторной онкологической помощи; ТСО — травматологическое стационарное отделение; ХСО — хирургическое стационарное отделение; ПМСП — первичная медико-санитарная помощь; ВОП — врач общей практики; ФАП — фельдшерско-акушерский пункт; ФП — фельдшерский пункт; СМП — скорая медицинская помощь.

Система оказания первичной медико-санитарной помощи

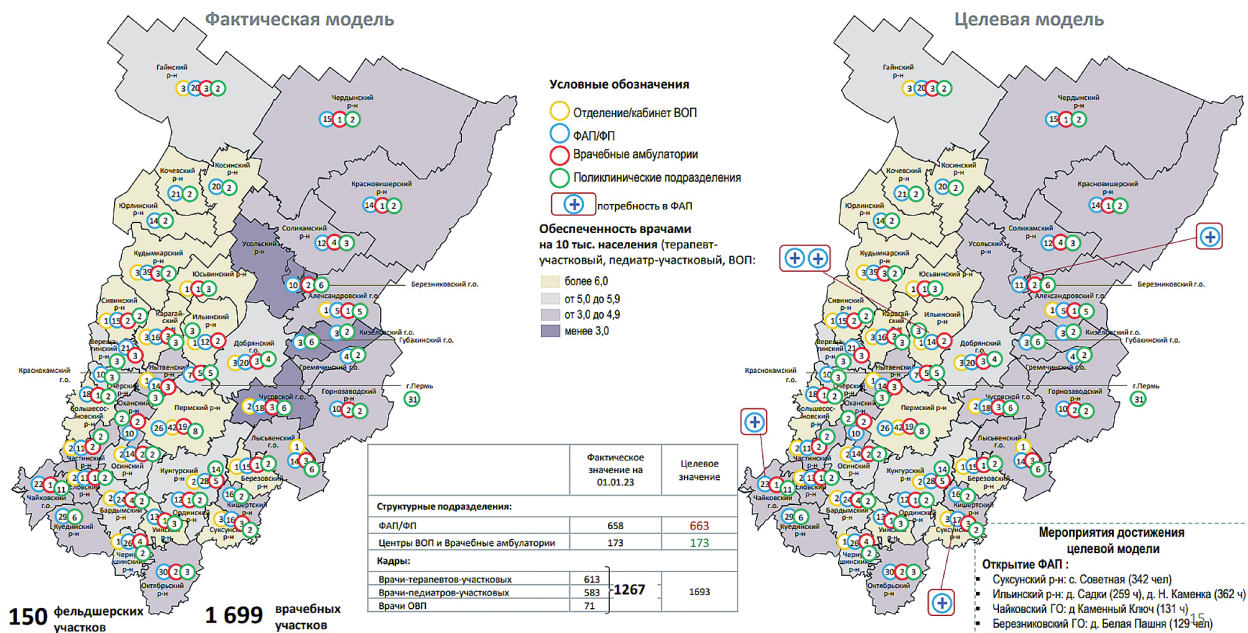


Рисунок 1 — Визуализация данных по фактической и целевой моделям системы оказания ПМСП в Пермском крае.

использования яркой видимой подписи на карте или использования графических элементов, иконок, соответствующих смыслу рассматриваемой карты.

Сформированные схемы территориального планирования содержали информацию о ресурсном обеспечении с учетом специфики оказания медицинской помощи по профилю (ключевые элементы системы здравоохранения региона, участвующие в оказании медицинской помощи в каждом районе). Для каждой из рассматриваемых систем оказания медицинской помощи соответствующие элементы приведены в табл. 1.

Предварительный анализ и оценка необходимых преобразований включала в себя отображение предлагаемых изменений на графической карте в виде целевого состояния в каждом из районов. Типовой перечень преобразований для достижения целевого состояния состоял из реализации следующих мероприятий:

- a. реорганизация объекта;
- b. добавление объекта;

- c. оснащение оборудованием;
- d. оснащение кадрами.

Дополнительно к сформированным схемам территориального планирования по каждому направлению приводился общий объем необходимых преобразований для достижения целевой модели.

Пример итоговой визуализации данных при территориальном планировании системы оказания первичной медико-санитарной помощи представлен на рис. 1.

Полученные материалы использовались в серии совещаний на уровне регионального органа управления здравоохранением, а также при проведении консультаций с профильными специалистами ЦНИИОИЗ и Минздрава России. В ходе обсуждений по каждой из систем был сформирован общий перечень необходимых преобразований и согласовано целевое состояние по каждой системе ПМСП на уровне региона. Набор картографических материалов был предоставлен ЛПР, отвечающим за достижение целевого состояния на уровне районов и региона в целом.

ОБСУЖДЕНИЕ

Был сформирован унифицированный подход к разработке визуальных картографических материалов в различных регионах. В случае, если пациенты госпитализируются по рассматриваемому профилю медицинской помощи преимущественно в экстренном состоянии, то характеристикой доступности расположения МО или их структурных подразделений может служить среднее время доезда скорой медицинской помощи, рассчитываемого по графу дорог. При наличии крупных водоемов и рек необходимо учитывать возможность доставки пациентов в экстренном состоянии по воде. В отсутствии наземного и водного транспортного сообщения рассматривается время доставки авиатранспортом.

Если рассматривается доступность госпитализации пациентов в плановом порядке, то время доезда до МО является второстепенным фактором, а на первый план выходит обеспеченность койками по профилю, достаточность которой можно оценить сравнением с общероссийским показателем и показателем расположенных рядом субъектов.

В комплексе с оценкой эффективности работы коечного фонда по профилям разработка схем территориального планирования позволяет не только определить число коек, не обеспеченных госпитализациями, но и ответить на вопрос о необходимом объеме сокращения коечного фонда в конкретных МО и их структурных подразделениях.

Разработка схем территориального планирования регионов позволяет определить проблемы доступности медицинской помощи, нерационального расположения объектов здравоохранения, приводящее как к необоснованному скоплению МО и их

подразделений, так и к увеличению времени доезда при госпитализации пациентов из районов, где отсутствуют отдельные ресурсы здравоохранения.

В ходе разработки были выявлены ограничения технического и организационного характера. Поскольку для выработки мероприятий требуется многофакторный анализ и знание региональной специфики, невозможно полностью автоматизировать процесс формирования перечня мероприятий по достижению целевого состояния в схемах территориального планирования.

Применение инструментов визуализации картографических материалов позволяет сократить время на восприятие табличных данных у ЛПР, снизить требования к компетенциям в обработке и визуализации геоданных. Это позволило в сжатые сроки выработать необходимые решения и обсудить перечень необходимых мероприятий для достижения целевого состояния.

ЗАКЛЮЧЕНИЕ

Для принятия взвешенных управленческих решений по вопросам организации здравоохранения, включая подготовку схем территориального планирования, необходимо осуществлять рациональный подход к выбору визуальных средств отображения результатов анализа данных.

Работа с инструментами визуализации позволяет ускорить процесс обсуждения и принятия управленческих решений. Для корректной интерпретации результатов визуализации данных и принятия рациональных решений у ЛПР должны быть сформированы базовые компетенции визуальной коммуникации и визуальной грамотности.

ЛИТЕРАТУРА/REFERENCES

1. Bloomberg Finance L.P. [updated 2022 May 5; cited 2023 August 17]. Available at: <https://www.bloomberg.com/press-releases/2022-05-05/study-reveals-massive-incentive-to-activate-unused-data>.
2. Гусев А.В., Реброва О.Ю. Осведомленность и мнения руководителей в сфере здравоохранения России о медицинских технологиях искусственного интеллекта // Врач и информационные технологии. — 2023. — №1. — С. 28-39. [Gusev AV, Rebrova OYu. Survey of awareness and opinions of Russian healthcare managers on artificial intelligence. Medical doctor and information technology. 2023; 1: 4: 28-39. (In Russ.)] doi: 110.25881/18110193_2023_1_28.

3. Кобякова О.С., Поликарпов А.В., Голубев Н.А., Рябков И.В., Лисненко А.А. Трансформация медицинской статистики в период пандемии новой коронавирусной инфекции (COVID-19) // Проблемы социальной гигиены, здравоохранения и истории медицины. — 2021. — №6 — С.1439-1445. [Kobyakova OS, Polikarpov AV, Golubev NA, Ryabkov IV, Lisnenko AA. The transformation of medical statistic during pandemic of new coronavirus infection (COVID-19). Problemy social'noj gigieny, zdravoohranenija i istorii mediciny. 2021; 29(6):1 439-1445 (In Russ.))] doi: 10.14341/probl201058491-103.
4. Zhao L, Ye W. Visualization as Infrastructure: China's Data Visualization Politics during COVID-19 and their Implications for Public Health Emergencies. *Convergence*. 2022; 28(1): 13-34. doi: 10.1177/13548565211069872.
5. Ronquillo JG, Lester WT, Zuckerman DM. Using informatics to guide public health policy during the COVID-19 pandemic in the USA. *J Public Health (Oxf)*. 2020; 42(4): 660-664. doi: 10.1093/pubmed/fdaa081.
6. Гусев А.В., Новицкий Р.Э. Технологии прогнозной аналитики в борьбе с пандемией COVID-19 // Врач и информационные технологии. — 2020. — №4. — С.24-33. [Gusev AV, Novitsky RE. Predictive analytics technologies in the management of the COVID-19 pandemic. Medical doctor and information technology. 2020; 4: 36-45. (In Russ.))] doi: 10.37690/1811-0193-2020-4-24-33
7. Мелик-Гусейнов Д.В. и др. Регрессионные модели прогнозирования количества летальных исходов при новой коронавирусной инфекции // Современные технологии в медицине. — 2020. — Т.12. — №2. — С.6-13. [Melik-Huseynov DV, et al. Regression Models Predicting the number of deaths from the new coronavirus infection. *Sovremennye tehnologii v medicine*. 2020; 12(2): 6-13. (In Russ.))] doi: 10.17691/stm2020.12.2.01.
8. Fortune Business Pvt. Ltd [updated 2023 March 2; cited 2023 August 17]. Available at: <https://www.fortunebusinessinsights.com/business-intelligence-bi-market-103742>.
9. Аббасов И.Б., Дешмух Р.Р. Применение искусственного интеллекта для медицинской визуализации // МНИЖ. — 2021. — №12-1(114). — С.43-49. [Abbasov IB, Deshmukh RR. Application of artificial intelligence for medical imaging. *International Research Journal*. 2021; 12-1(114): 43-49. (In Russ.))] doi: <https://doi.org/10.23670/IRJ.2021.114.12.005>.
10. Гусев А.В., Морозов С.П., Кутичев В.А., Новицкий Р.Э. Нормативно-правовое регулирование программного обеспечения для здравоохранения, созданного с применением технологий искусственного интеллекта, в Российской Федерации // Медицинские технологии. Оценка и выбор. — 2021. — №43(1) — С.36-45. [Gusev AV, Morozov SP, Kutichev VA, Novitsky RE. Legal regulation of artificial intelligence software in healthcare in the Russian Federation. *Medical Technologies. Assessment and Choice*. 2021; 43(1): 36-45. (In Russ.))] doi: <https://doi.org/10.17116/medtech20214301136>.
11. Ившин А.А., Багаудин Т.З., Гусев А.В. Искусственный интеллект на страже репродуктивного здоровья // Акушерство и гинекология. — 2021. — №5. — С.17-24. [Ivshin AA, Bagaudin TZ, Gusev AV. Artificial intelligence on guard of reproductive health. *Akusherstvo i Ginekologiya*. 2021; 5: 17-24. (In Russ.))] doi: <https://dx.doi.org/10.18565/aig.2021.5.17-24>.
12. Интернет-портал «Российской газеты» [доступ от 17.08.2023]. Доступно по: <https://rg.ru/2021/06/02/reg-szfo/skorost-priniatiia-reshenij-v-zdravoohranenii-vyrosla-v-razy.html>. [Internet-portal «Rossijskoj gazety». Available at: <https://rg.ru/2021/06/02/reg-szfo/skorost-priniatiia-reshenij-v-zdravoohranenii-vyrosla-v-razy.html>. cited 2023 August 17. (In Russ.))]
13. Голубчиков Ю.Н., Тикунов В.С., Тикунова И.Н. Геоинформационная система для демографических исследований // Географический вестник. — 2014. — №3(30). — С.130-137. [Golubchikov YuN, Tikunov VS, Tikunova IN. Geoinformation system for demographic research. *Geograficheskij vestnik*. 2014; 3(30): 130-137. (In Russ.))]

14. Кокорина И.П. К вопросу о применении ГИС-технологий в медико-географическом картографировании // Интерэкспо Гео-Сибирь. — 2006. — Т.1. — №2. — С.252-255. [Kokorina IP. K voprosu o primenenii GIS-tehnologij v mediko-geograficheskom kartografirovanii. Interjekspo Geo-Sibir'. 2006; 2: 252-255. (In Russ.)]
15. Поликарпов А.В., Голубев Н.А., Левахина Ю.С., Махов В.А. Геоинформационная система как элемент территориального планирования в здравоохранении // Вестник Росздравнадзора. — 2022. — №6. — С.68-74. [Polikarpov AV, Golubev NA, Levakhina YuS, Makhov VA. Geoinformation system as an element of territorial planning in health cares. Vestnik Roszdravnadzora. 2022; 6: 68-74. (In Russ.)]
16. Указ Президента РФ от 6 июня 2019 г. №254 «О Стратегии развития здравоохранения в Российской Федерации на период до 2025 года». Доступно по: <https://base.garant.ru/72264534/>. Ссылка активна на 01.09.2023. [Decree of the President of the Russian Federation №254 of 6 June 2019. «O Strategii razvitiya zdravoochraneniya v Rossijskoj Federacii na period do 2025 goda». Available at: <https://base.garant.ru/72264534/>. cited 2023 Sep. 1. (In Russ.)]
17. Поликарпов А.В., Огрызко Е.В., Курсекова И.В. Анализ нормативного регулирования территориального планирования в здравоохранении // Современные проблемы здравоохранения и медицинской статистики. — 2022. — №4. — С.610-628. [Polikarpov AV, Ogryzko EV, Kursekova IV. Analysis of the regulatory regulation of territorial planning in healthcare. 2022; 4: 610-628. (In Russ.)] doi: 10.24412/2312-2935-2022-4-610-628.
18. The Problem of Abortion and the Doctrine of the Double Effect Phillipa Foot/ Foot Ph. The Problem of Abortion and the Doctrine of the Double Effect. Oxford Review. 1967; 5: 5-15.
19. Афанасьев А.А. Технология визуализации данных как инструмент совершенствования процесса поддержки принятия решений // Инженерный вестник Дона. — 2014. — Т.31. — №4-1. — С.60. [Afanas'ev A. A. Tekhnologiya vizualizacii dannyh kak instrument sovershenstvovaniya processa podderzhki prinyatiya reshenij. Inzhenernyj vestnik Dona. 2014; 31(4-1): 60. (In Russ.)]
20. Grincheva N. Beyond the scorecard diplomacy: From soft power rankings to critical inductive geography. Convergence. 2022; 28(1): 70-91. doi: 10.1177/13548565221079158.
21. Amit-Danhi, Eedan R. Facts of (financial) life: Political power and Re-visualization strategies. Convergence. 2022; 28(1): 92-108. doi: 10.1177/13548565211045533.
22. Постановление Правительства РФ №2497 от 29.12.2022 «О Программе государственных гарантий бесплатного оказания гражданам медицинской помощи на 2023 год и на плановый период 2024 и 2025 годов». Доступно по: <http://static.government.ru/media/files/FQATIOfojXIUYX8cwl2X7ugkeKRrRGjb.pdf>. Ссылка активна на 01.09.2023. [Resolution of the Government of the Russian Federation №2497 of 29 December 2022. «O Programme gosudarstvennyh garantij besplatnogo okazaniya grazhdanam medicinskoj pomoshhi na 2023 god i na planovyj period 2024 i 2025 godov». Available at: <http://static.government.ru/media/files/FQATIOfojXIUYX8cwl2X7ugkeKRrRGjb.pdf>. cited 2023 Sep. 1. (In Russ.)]

23. Постановление Правительства Российской Федерации №178 от 24.03.2007 «Об утверждении Положения о согласовании проектов схем территориального планирования субъектов Российской Федерации». Доступно по: <https://base.garant.ru/12152642/>. Ссылка активна на 01.09.2023. [Resolution of the Government of the Russian Federation №178 of 24 March 2007. «Ob utverzhdenii Polozhenija o soglasovanii proektov shem territorial'nogo planirovanija sub#ektov Rossijskoj Federacii». Available at: <https://base.garant.ru/12152642/>. cited 2023 Sep. 1. (In Russ).]
24. Приказ Минстроя России №1034/пр от 30 декабря 2016 «Об утверждении СП 42.13330 «СНиП 2.07.01-89* Градостроительство. Планировка и застройка городских и сельских поселений». Доступно по: <https://minstroyrf.gov.ru/docs/14465/>. Ссылка активна на 01.09.2023. [Resolution of the Government of the Russian Federation №1034 of 30 December 2016. «Ob utverzhdenii SP 42.13330 «SNiP 2.07.01-89* Gradostroitel'stvo. Planirovka i zastrojka gorodskih i sel'skih poselenij». Available at: <https://minstroyrf.gov.ru/docs/14465/>. cited 2023 Sep. 1. (In Russ).]
25. Постановление Главного государственного санитарного врача РФ №44 от 24 декабря 2020 г. «Об утверждении санитарных правил СП 2.1.3678-20 «Санитарно-эпидемиологические требования к эксплуатации помещений, зданий, сооружений, оборудования и транспорта, а также условиям деятельности хозяйствующих субъектов, осуществляющих продажу товаров, выполнение работ или оказание услуг». Доступно по: <https://www.garant.ru/products/ipo/prime/doc/400063274/>. Ссылка активна на 01.09.2023. [Resolution of the Chief State Sanitary Doctor of the Russian Federation №44 of 24 December 2020. «Ob utverzhdenii sanitarnyh pravil SP 2.1.3678-20 «Sanitarno-jepidemiologicheskie trebovanija k jekspluatacii pomeshhenij, zdanij, sooruzhenij, oborudovanija i transporta, a takzhe uslovijam dejatel'nosti hozjajstvujushhih sub#ektov, osushhestvljajushhih prodazhu tovarov, vypolnenie rabot ili okazanie uslug». Available at: <https://www.garant.ru/products/ipo/prime/doc/400063274/>. cited 2023 Sep. 1. (In Russ).]

ПОРЕЦКОВА Г.Ю.,

д.м.н., доцент, ФГБОУ ВО СамГМУ Минздрава России, г. Самара, Россия, e-mail: g.yu.poreckova@samsmu.ru

ТЯЖЕВА А.А.,

к.м.н., ФГБОУ ВО СамГМУ Минздрава России, г. Самара, Россия, e-mail: a.a.tyazheva@samsmu.ru

ПЛАХОТНИКОВА С.В.,

к.м.н., ФГБОУ ВО СамГМУ Минздрава России, г. Самара, Россия, e-mail: s.v.plahotnikova@samsmu.ru

ЧАПЛЫГИН С.С.,

к.м.н., ФГБОУ ВО СамГМУ Минздрава России, г. Самара, Россия, e-mail: s.s.chaplygin@samsmu.ru

НАПАЛКОВА С.А.,

ФГБОУ ВО СамГМУ Минздрава России, г. Самара, Россия, e-mail: s.a.napalkova@samsmu.ru

БАРИНОВ И.В.,

к.м.н., ФГБОУ ВО СамГМУ Минздрава России, г. Самара, Россия, e-mail: i.v.barinov@samsmu.ru

ПАНОВА Н.И.,

ФГБОУ ВО СамГМУ Минздрава России, г. Самара, Россия, e-mail: n.i.panova@samsmu.ru

ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ ЦИФРОВЫХ ТЕХНОЛОГИЙ ПРИ ПРОФИЛАКТИЧЕСКОМ ОСМОТРЕ ШКОЛЬНИКОВ Г. САМАРЫ

DOI: 10.25881/18110193_2023_3_84

Аннотация. Одной из ведущих задач здравоохранения является сохранение здоровья детского населения, что достигается путем проведения комплексной оценки состояния здоровья. Необходимо своевременно выявлять отклонения в состоянии здоровья детей при проведении периодических медицинских осмотров, что обосновывает разработку более точных методов, возможность извлечения полезной информации по данным, полученным в ходе комплексного осмотра детей.

Цель работы: усовершенствовать процедуру проведения профилактических осмотров детей школьного возраста за счёт использования цифровых технологий и дистанционной передачи данных.

Материалы и методы: с использованием цифровых технологий проведён профилактический осмотр 128 детей в возрасте 14-17 лет г. Самары, в группу сравнения вошли 54 подростка, которым профилактический осмотр проводился с использованием традиционных методов; группы сопоставимы по полу и среднему возрасту.

Результаты: в результате работы было показано, что использование цифровых технологий позволяет сократить работу медицинского персонала при проведении медицинского осмотра школьников, а также с использованием аналитического программно-аппаратного комплекса было выявлено 47,8% детей, имеющих нарушения физического развития, 3,8% детей со впервые выявленным нарушением бронхиальной проводимости и 4,6% детей с артериальной гипертензией.

Выводы: использование цифровых технологий позволяет снизить количество возможных ошибок при фиксации антропометрических и физиометрических показателей и повысить оперативность принятия врачебных решений при проведении профилактических осмотров детей.

Ключевые слова: цифровые технологии, профилактические осмотры, комплексная оценка состояния здоровья, дети, подростки.

Для цитирования: Порецкова Г.Ю., Тяжева А.А., Плахотникова С.В., Чаплыгин С.С., Напалкова С.А., Баринов И.В., Панова Н.И. Возможности использования цифровых технологий при профилактическом осмотре школьников г. Самары. *Врач и информационные технологии.* 2023; 3: 84-92. doi: 10.25881/18110193_2023_3_84.

PORETSKOVA G.YU.,

DSc, Associate Professor, Samara State Medical University, Samara, Russia, e-mail: g.yu.poreckova@samsmu.ru

TYAZHEVA A.A.,

PhD, Samara State Medical University, Samara, Russia, e-mail: a.a.tyazheva@samsmu.ru

PLAKHOTNIKOVA S.V.,

PhD, Samara State Medical University, Samara, Russia, e-mail: s.v.plahotnikova@samsmu.ru

CHAPLYGIN S.S.,

PhD, Samara State Medical University, Samara, Russia, e-mail: s.s.chaplygin@samsmu.ru

NAPALKOVA S.A.,

Samara State Medical University, Samara, Russia, e-mail: s.a.napalkova@samsmu.ru

BARINOV I.V.,

PhD, Samara State Medical University, Samara, Russia, e-mail: i.v.barinov@samsmu.ru

PANOVA N.I.,

Samara State Medical University, Samara, Russia, e-mail: n.i.panova@samsmu.ru

POSSIBLE APPLICATIONS OF DIGITAL TECHNOLOGIES IN THE PREVENTIVE EXAMINATION OF SCHOOLCHILDREN IN SAMARA

DOI: 10.25881/18110193_2023_3_84

Abstract. *One of the key tasks of healthcare is to preserve the health of the child population, which is achieved through comprehensive health assessment. Timely identification of negative changes in children health during periodic medical examinations is of great importance. This justifies the development of more accurate methods to extract useful information from the data obtained during a comprehensive examination of children.*

Aim: to improve the procedure for preventive examinations of school-age children using digital technologies and remote data transmission.

Methods: 128 children aged 14-17 years living in Samara were examined using digital technologies. The comparison group included 54 adolescents of the similar gender and average age distribution within the group.

Results: the use of information technology allows reducing the work load of medical personnel during a medical examination of schoolchildren. Analytical software and hardware complex identified that 47.8% of children had physical development disorders, 3.8% of children had newly diagnosed bronchial conduction disorders and 4.6% of children had arterial hypertension.

Conclusions: the use of digital technologies allows reducing the number of possible errors when fixing anthropometric and physiometric data and increasing the efficiency of medical decision-making during preventive examinations of children.

Keywords: *digital technologies, preventive inspection, comprehensive health assessment, children, adolescents.*

For citation: *Poretskova G.Yu., Tyazheva A.A., Plakhotnikova S.V., Chaplygin S.S., Napalkova S.A., Barinov I.V., Panova N.I. Possible applications of digital technologies in the preventive examination of schoolchildren in Samara. Medical doctor and information technology. 2023; 3: 84-92. doi: 10.25881/18110193_2023_3_84.*

Сохранение здоровья детского населения является важнейшей задачей не только здравоохранения, но и государственной политики в целом. Решение этой важнейшей задачи в нашей стране достигается путем функционирования системы комплексной оценки состояния здоровья несовершеннолетних, принципы которой разработаны и поддерживаются ведущими учеными «Союза Педиатров России» и НИИ гигиены детей и подростков [1, 2]. На государственном уровне данная программа закреплена принятием ряда нормативных документов, в частности приказа Минздрава России от 10.08.2017 № 514н «О Порядке прохождения несовершеннолетними медицинских осмотров, в том числе при поступлении в образовательные учреждения и в период обучения в них» [3].

Для своевременного выявления отклонений в состоянии здоровья детей при проведении периодических медицинских осмотров возрастает необходимость извлечения полезной информации по данным, полученным в ходе комплексного осмотра детей, и требуется развитие методологии анализа информации в этой области.

В условиях взрывного роста медицинской информации и количества профилактических, диагностических и лечебных технологий врач должен потратить много времени и иметь специальные навыки для поиска, анализа и применения этой информации на практике. Кроме этого, обеспеченность обучающихся врачами и средним медицинским персоналом в дошкольных и общеобразовательных организациях явно недостаточна, а в учреждениях начального и среднего профессионального образования катастрофична [2, 4]. Все это зачастую может приводить к снижению качества медицинских осмотров детей и подростков.

В последние годы в здравоохранении активно внедряются процессы «цифровой трансформации». В настоящее время опубликовано достаточно научных работ, описывающих успехи в области цифровых технологий в здравоохранении [5, 6], внедрения в клиническую практику технологий «цифровой медицины», совершенствования алгоритмов диагностики, лечения и применения средств автоматизации [6, 7].

Цифровые врачебные кейсы разработаны для мониторинга пациентов с социально-значимыми заболеваниями, что позволяет

дистанционно мониторировать состояние больных и, при необходимости, производить коррекцию лечения [8–10].

Кроме того, рутинные осмотры пациентов, особенно при массовом скрининге, имеют ряд факторов, снижающих достоверность исследования, — самостоятельное использование бесконтактных термометров, измерение показателей артериального давления, когда медицинская сестра или сам пациент фиксирует усредненные показатели [11].

Цифровая трансформация процессов проведения медицинских осмотров детей позволяет системе здравоохранения выйти на принципиально новый качественный уровень [6]. Прежде всего потому, что применение цифровых технологий при комплексной оценке состояния здоровья школьников объективизирует процесс сбора и фиксации информации, при этом минимизируется влияние человеческого фактора, сокращается время, затраченное на проведение комплексного обследования и осмотра ребенка, кроме того врачу не придется затрачивать время на анализ полученных данных, поскольку он происходит в автоматизированном режиме, тем самым выстраивая эффективную модель управления рисками — своевременное выявление отклонений в состоянии здоровья и назначение превентивных профилактических мер [6,12].

Цель исследования: усовершенствовать процедуру проведения профилактических осмотров детей школьного возраста в г. Самаре, повысить их качество за счёт создания аналитического программно-аппаратного комплекса (АПАК) с перспективой интеграции в систему обмена информацией в здравоохранении.

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Для реализации поставленной цели был разработан АПАК для оценки состояния здоровья детей различного возраста. Комплекс представляет собой систему, состоящую из мобильного приложения для медицинского работника, веб-приложения для врача, кейса с портативными приборами для измерения физиологических показателей пациентов с возможностью беспроводной передачи данных, а также возможностью ручного ввода результатов измерений.

В ходе исследования нами было обследовано с помощью комплекса АПАК 128 детей (основная

Таблица 1 — Комплектация кейса

Наименование прибора	Измеряемый параметр	Наименование в интерфейсе	Единицы измерения	Количество знаков	Способ внесения показаний в АПАК
Пульсоксиметр	частота пульса	Пульс	Ударов/мин	XX	автоматически
	процент сатурации кислорода	SO2	%	XX	автоматически
Термометр	температура тела	Температура	градус Цельсия	XX.X	автоматически
Весы	вес	Вес	кг	XX.X	автоматически
Ростомер	рост	Рост, см	см	XXX.X	Вручную
Тонومتر	артериальное давление	АД сист	мм рт. ст.	XX	автоматически
		АД диаст	мм рт. ст.	XX	автоматически
Аппарат ЭКГ	электрические потенциалы сердца	ЭКГ	Частота сердечных сокращений	XX	автоматически
			График ЭКГ	рисунок	автоматически
Спирометр (используется для детей от 5 лет)		фЖЕЛ	Литр	X.XX	автоматически
		ОФВ1		X.XX	автоматически
		ПОСвыд		X.XX	автоматически
		ОФВ1/ фЖЕЛ		XX.X	автоматически
		МОС25		X.XX	автоматически
		МОС 50		X.XX	автоматически
		МОС75		X.XX	автоматически
СОС25-75		X.XX	автоматически		

группа), обучающихся в образовательных организациях (школы) г. Самары в возрасте 14–17 лет (мальчиков 54, средний возраст $15,6 \pm 0,16$ лет, девочек 74, средний возраст $15,7 \pm 0,14$ лет). Группу сравнения составили 54 ребенка (мальчиков 31, средний возраст $15,1 \pm 0,11$ лет, девочек 23, средний возраст $15,9 \pm 0,13$ лет), обследование которых в рамках профилактического осмотра проводилось рутинным методом. Параметры физического развития оценивались у 113 детей основной группы и у 54 детей контрольной группы.

Методы исследования: антропометрические методы оценки физического развития (изменение массы, роста, расчёт индекса массы тела (m/h^2 , где m — масса тела, h^2 — рост в метрах в квадрате)), оценка функционального состояния

— определение жизненной емкости легких, спирографии с автоматическим расчетом индексов (Тиффно = $ОФВ1 / фЖЕЛ$, где $ОФВ1$ — объем форсированного выдоха за первую секунду, $фЖЕЛ$ — форсированная жизненная емкость), измерение уровня артериального давления. В исследовании применялись количественные методы измерения и статистические методы обработки полученной информации (расчет доверительного интервала, отношения шансов (ОШ) и критерия χ^2 -критерия).

В комплектацию кейса для измерения физиологических показателей входили медицинские приборы, имеющие Bluetooth функцию, регистрационное удостоверение и выполняющие виды измерений, отраженные в таблице 1.

При проведении стандартного профилактического осмотра на этапе доврачебного осмотра медицинская сестра (или фельдшер) измеряет показатели физического развития ребенка и фиксирует их в карте профилактического осмотра, затем на конечном этапе при проведении комплексной оценки состояния здоровья педиатр вручную оценивает массо-ростовые показатели на основании имеющихся стандартов.

При проведении профилактического осмотра рутинным методом контрольной группы нами было рассчитано, что среднее затраченное время на осмотр с фиксацией показателей одного ребенка составляет 3,5 минуты, при использовании системы АПАК на осмотр затрачивается 2,25 минуты с автоматизированной фиксацией результатов.

При оценке физического развития методом сигмальных отклонений (SD) по индексу массы тела (ИМТ) — величина z-score (ВОЗ) — «средний уровень» показателя констатируется, если сигмальное отклонение фактической массы тела от среднеарифметического значения не превышает ± 1 SD; «ниже среднего» — от -1 до -2 SD; «низкий» — менее -2 SD; «выше среднего» — от +1 до +2 SD; «высокий» — более +2 SD [13].

Согласно Федеральным клиническим рекомендациям по диагностике и лечению ожирения у детей и подростков и рекомендациям Российской ассоциации эндокринологов (2014 г.) недостаточность питания характеризуется ИМТ < -2,0 SD; избыточная масса тела — от +1,0 SD до +2,0 SD; ожирение — > +2,0 SD [14].

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Благодаря использованию АПАК, параметры физического развития, внесённые автоматически через Bluetooth, вручную и рассчитанные,

оцениваются системой автоматически уже на этапе доврачебного осмотра, при этом сокращается время, затраченное врачом, и появляется возможность выявить детей с дисгармоничным развитием различной степени.

Так, согласно нашим данным, при оценке физического развития методом сигмальных отклонений по ИМТ среди обследованных детей 47,8% (55 детей) имеют нарушение физического развития, при этом только 12,7% имеют выраженные нарушения (недостаточность питания — 4,3%, ожирение — 7,8%), а 35,1% подростков имеют пограничные нарушения в виде избыточной или низкой массы тела (табл. 2).

Согласно региональным данным, при проведении медицинского осмотра школьников г. Самары без использования цифровых технологий нарушения физического развития выявляются значительно реже — недостаточность питания 1,9%, ожирение 6,4% [15]. Несвоевременная диагностика нарушений параметров физического развития влечет за собой рост таких социально-значимых заболеваний как метаболический синдром и сердечно-сосудистые заболевания [16].

Полученные данные показывают, что использование системы АПАК позволяет своевременно выявлять и более точно определять степени нарушения физического развития, при этом облегчая врачу-педиатру задачу оценки большого количества антропометрических параметров при массовом проведении профилактического осмотра школьников.

К физиометрическим показателям физического развития детей относятся жизненная емкость легких (ЖЕЛ) и показатели уровня артериального давления (систолическое и диастолическое артериальное давление).

При проведении рутинного медицинского осмотра на доврачебном этапе измерение ЖЕЛ

Таблица 2 — Отклонения в показателях физического развития по индексу массы тела, определяемые с помощью АПАК

Пол	Очень низкие показатели (n = 5)	Низкие показатели (n = 20)	Высокие показатели (n = 21)	Очень высокие показатели (n = 9)
Мальчики (n = 48)	2 (4,2%)	9 (18,7%)	8 (16,7%)	7 (14,6%)
Девочки (n = 65)	3 (4,6%)	11 (16,9%)	13 (20%)	2 (3,1%)

Таблица 3 — Показатели жизненной емкости легких детей-подростков при проведении медицинского осмотра с использованием системы АПАК

	Результаты при использовании АПАК	Результаты рутинного осмотра	χ^2
Неудовлетворительные показатели	5 (3,8%)	17 (31,5%)	$\chi^2 = 19,6$ (при критическом значении 7,66)
Удовлетворительные показатели	9 (6,9%)	0	$\chi^2 = 3,7$ (при критическом значении 2,54)
Хорошие показатели	11 (8,5%)	0	$\chi^2 = 4,5$ (при критическом значении 3,08)
Отличные показатели	103 (79,8%)	37 (68,5%)	$\chi^2 = 0,4$ (при критическом значении 39,57)

проводит медицинская сестра, записывая результаты в карту каждого ребенка, далее при проведении комплексной оценки врачу приходится, используя центильные таблицы, интерпретировать результаты измерений, что требует большого количества времени и, в конечном итоге может привести к погрешностям оценки результатов. В систему АПАК включен автоматический анализ данного показателя индивидуально для каждого пациента, и при формировании итогового отчета исследования врач-педиатр сразу видит результат, тем самым уменьшается вероятность ошибки (табл. 3).

Отличные показатели жизненной емкости легких оцениваются не ниже, чем по 6 центильному интервалу, хорошие — не ниже, чем по 4–5 центильным интервалам, удовлетворительные — не ниже, чем по 2–3 центильным интервалам, неудовлетворительные — по 1 центильному интервалу. Высокие значения жизненной емкости легких всегда указывают на хороший уровень физической способности школьника.

Значения ЖЕЛ учитываются для оценки функционального состояния дыхательной системы, однако недостаточно использовать только этот параметр. Проведение процедуры спирометрии не входит в определённый приказом перечень исследований при профилактическом осмотре детей-подростков. В автоматическом комплексе АПАК имеется спирометр, что позволяет расширить оценку функционального состояния органов дыхания. Дополнительное проведение

спирометрии не приводило к увеличению времени медицинского осмотра, поскольку показатели снимались одновременно при проведении исследования ЖЕЛ.

Одним из важных показателей, отражающих нарушения бронхиальной проводимости является соотношение объема форсированного выдоха в 1 минуту к ЖЕЛ — индекс Тиффно. Благодаря спирометрии удалось выявить 5 человек (3,8%) с признаками нарушения бронхиальной проводимости — 3 ребенка (2,3%) с умеренными отклонениями и 2 подростка (1,5%) с выраженными отклонениями индекса Тиффно. Все эти дети были направлены на консультацию к пульмонологу.

Измерение артериального давления (АД) при профилактических осмотрах также проводится на доврачебном этапе. При этом оценка показателей для врача является ресурсозатратной процедурой, поскольку нормативы систолического (САД) и диастолического (ДАД) артериального давления определяются центильным методом с обязательной оценкой на первом этапе перцентилья роста ребенка.

Система АПАК упрощает работу врача-педиатра, поскольку после измерения АД расчет результата производится автоматически на основании нормативов с учётом пола и возраста пациента и врач получает заключение по данному параметру при формировании индивидуально и комплексного итогового отчёта по осмотру.

При проведении медицинского осмотра с использованием АПАК было установлено, что

Таблица 4 — Результаты показателей артериального давления при проведении медицинского осмотра подростков с использованием АПАК

	Артериальная гипертензия		Высокое нормальное АД		Артериальная гипотензия	
	САД	ДАД	САД	ДАД	САД	ДАД
Девочки	0	1	2	0	32	13
Мальчики	6	1	3	3	9	12
Всего	6	2	5	3	41	25

у 6 детей выявлен синдром артериальной гипертензии (4,6%), артериальная гипотония определена у 16 детей (12,4%), в то время как при рутинном медицинском осмотре выявляемость данной патологии неоднозначна — артериальная гипертензия выявлялась чаще (12 детей — 22,2%), при этом при повторном измерении у более половины детей АД в пределах возрастной нормы, а артериальная гипотония выявлялась значительно реже — 3 ребенка (5,6%) — ОШ для гипертензии 0,42 (ДИ 0,13–1,36), ОШ для артериальной гипотензии -2,25 (ДИ 0,63–8,03). В то же время часть детей имели показатели высокого нормального артериального давления (АД-САД и/или ДАД, уровень которого > или равен 90-му и <95-го перцентеля кривой распределения АД в популяции для соответствующего возраста, пола или роста), что требует пристального медицинского контроля за этой группой пациентов (табл. 4). При профилактическом медицинском осмотре без использования цифровых технологий, данная группа детей не выделяется.

ОБСУЖДЕНИЕ

В настоящее время уже внедренные программные комплексы в систему здравоохранения РФ позволяют в основном проводить анализ полученных данных, автоматизировать выделение групп здоровья, формировать отчетность, в том числе по проводимым профилактическим осмотрам школьников [17–19], при этом рутинное обследование пациента по-прежнему остается в рамках деятельности медицинского персонала. В ходе нашего исследования было показано, что автоматизированные поликлинические комплексы с использованием

стандартных приборов для проведения медицинского осмотра и автоматизированной передачи данных позволяют минимизировать погрешности ручной обработки информации. При этом своевременно выявляются даже незначительные отклонения в параметрах физического развития ребенка, своевременно выявляются изменения при физиометрических исследованиях, данная информация находит отражение в автоматически сформированном отчете проведенного профилактического осмотра. В таком случае, значительно уменьшается нагрузка на врача-педиатра и повышается качество проводимых профилактических медицинских осмотров детей.

ВЫВОДЫ

Таким образом, использование автоматических поликлинических комплексов с функциями передачи данных позволяют повысить оперативность принятия врачебных решений при проведении профилактических осмотров детей, так как обеспечивается полнота и корректность данных, полученных в ходе осмотра, снижается количество возможных ошибок при фиксации и оценке антропометрических и физиометрических показателей. При этом облегчается работа врача при проведении массового профилактического осмотра школьников за счет автоматической фиксации, интерпретации полученных данных и их сохранения в индивидуальной карте каждого ребёнка и в сформированном итоговом отчете. Внедрение цифровых технологий в рамках автоматизированных систем позволяет решать определённый круг задач, что значительно улучшает качество работы врача.

ЛИТЕРАТУРА/REFERENCES

1. Рост и развитие детей. Контроль за состоянием здоровья детей в образовательных учреждениях. Руководство по амбулаторно-поликлинической педиатрии. / Под ред. А.А. Баранова. — 2-е изд., испр. и доп., 2009. — С.3-37 [Growth and development of children. Monitoring the health status of children in educational institutions. Guidelines for outpatient pediatrics. In: Rukovodstvo po ambulatorno-poliklinicheskoj pediatrii. A.A. Baranov, editor. 2-e izd., ispr. i dop. 2009: 3-37. (In Russ.)]
2. Кучма В.Р., Макарова Ю.А., Рапопорт И.К. Состояние медицинского обеспечения детей в образовательных организациях. // Вопросы школьной и университетской медицины и здоровья. — 2017. — №2. — С.37-45. [Kuchma VR, Makarova JuA, Rapoport IK. The state of medical support for children in educational institutions. Voprosy shkol'noj i universitetskoj mediciny i zdorov'ja. 2017; 2: 37-45. (In Russ.)]
3. Приказ МЗ РФ от 10.08.2017 №514н «О Порядке прохождения несовершеннолетними медицинских осмотров, в том числе при поступлении в образовательные учреждения и в период обучения в них». [Federal Law of Russian Federation 10/08/2017 «O Porjadke prohozhdenija nesovershennoletnimi medicinskih osmotrov, v tom chisle pri postuplenii v obrazovatel'nye uchrezhdenija i v period obuchenija v nih». (In Russ.)]
4. Кучма В.Р., Соколова С.Б., Рапопорт И.К., Макарова А.Ю. Организация профилактической работы в образовательных учреждениях: проблемы и пути решения // Гигиена и санитария. — 2015. — Т.94. — №1. — С.5-8. [Kuchma VR, Sokolova SB, Rapoport IK, Makarova AJu. Organizacija profilakticheskoj raboty v obrazovatel'nyh uchrezhdenijah: problemy i puti reshenija. Gigiena i sanitarija. 2015; 94(1): 5-8. (In Russ.)]
5. Бацина Е.А., Попсуйко А.Н., Артамонова Г.В. Цифровизация здравоохранения РФ: миф или реальность? // Врач и информационные технологии. — 2020. — №3. — С.73-80. [Bacina EA, Popsujko AN, Artamonova GV. Cifrovizacija zdravoohranenija RF: mif ili real'nost'? Vrach i informacionnye tehnologii. 2020; 3: 73-80. (In Russ.)] doi: 10.37690/1811-0193-2020-3-73-80.
6. Указ Президента Российской Федерации от 09.05.2017 №203 «О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы». Доступно по: <http://www.consultant.ru>. Ссылка активна на 17.07.2021. [Ukaz Prezidenta Rossijskoj Federacii ot 09.05.2017 №203 «O Strategii razvitija informacionnogo obshhestva v Rossijskoj Federacii na 2017–2030 gody». Available at: <http://www.consultant.ru>. Accessed 17.07.2021. (In Russ.)]
7. Макаров В.В., Асеев А.А., Крутелев К.Н. Применение телемедицины для совершенствования бизнес-процессов диагностики и лечения // Modern science and technology: Сборник статей II Международной научно-практической конференции, Петрозаводск, 20 февраля 2020 года. — Петрозаводск: Международный центр научного партнерства «Новая Наука», 2020: 45-50. [Makarov VV, Aseev AA, Krutelev KN. Primenenie telemediciny dlja sovershenstvovanija biznes-processov diagnostiki i lechenija. Modern science and technology: Sbornik statej II Mezhdunarodnoj nauchno-prakticheskoj konferencii, Petrozavodsk, 20 fevralja2020goda.Petrozavodsk:Mezhdunarodnyjcentrnauchnogopartnerstva«NovajaNauka»,2020:45-50. (In Russ.)]
8. Шаронова Л. А. Применение цифровых технологий в курации пациента с сахарным диабетом 2 типа // Опорный образовательный центр: Учебное пособие: сборник кейсов за 2021 год по развитию цифровых компетенций обучающихся по программам среднего профессионального и высшего образования. Том 3. — Казань: Университет Иннополис, 2021. — С.25-28. [Haronova LA. Primenenie cifrovyh tekhnologij v kuracii pacienta s saharnym diabetom 2 tipa. Opornyj obrazovatel'nyj centr : Uchebnoe posobie: sbornik kejsov za 2021 god po razvitiyu cifrovyh kompetencij obuchayushchihhsya po programmam srednego professional'nogo i vysshego obrazovaniya. Tom 3. Kazan': Universitet Innopolis, 2021. S.25-28. (In Russ.)]
9. Riaz MS, Atreja A. Personalized Technologies in Chronic Gastrointestinal Disorders: Self-monitoring and Remote Sensor Technologies. Clin Gastroenterol Hepatol. 2016; 14(12): 1697-1705. doi: 10.1016/j.cgh.2016.05.009.
10. Shinbane JS, Saxon LA. Virtual medicine: Utilization of the advanced cardiac imaging patient avatar for procedural planning and facilitation. J Cardiovasc Comput Tomogr. 2018; 12(1): 16-27. doi: 10.1016/j.jcct.2017.11.004.

11. Клепалова, Ю.И. Цифровизация медицинских осмотров работников транспорта // Цифровые технологии и право: Сборник научных трудов I Международной научно-практической конференции. В 6 т. Казань, 23 сентября 2022 года / Под редакцией И.П. Бегишева и др. Том 3. — Казань: Познание, 2022. — С.143-148. [KlepaloVA YUI. Cifrovizaciya medicinskih osmotrov rabotnikov transporta. Cifrovyte tekhnologii i pravo : Sbornik nauchnyh trudov I Mezhdunarodnoj nauchno-prakticheskoj konferencii. V 6 t. Kazan', 23 sentyabrya 2022 goda I. Tom 3. Kazan': Poznanie, 2022. S.143-148. (In Russ.)]
12. Свидетельство о государственной регистрации программы для ЭВМ №2019619657 Российская Федерация. Персонализированный подход к оценке и коррекции фактического питания лиц с низкой физической активностью: №2019618345: опублик. 22.07.2019 / Г. Ю. Порецкова, Д. В. Печуров, О. О. Биденко и др. [Svidetel'stvo o gosudarstvennoj registracii programmy dlja JeVM №2019619657 Rossijskaja Federacija. Personificirovannyj podhod k ocenke i korrekcii fakticheskogo pitaniya lic s nizkoj fizicheskoj aktivnost'ju : №2019618345: opubl. 22.07.2019. (In Russ.)]
13. Стандарты физического развития детей. ВОЗ. Доступно по: <https://www.who.int/tools/child-growth-standards/standards/weight-for-age>. Ссылка активна на 08.07.2023. [Standards of physical development of children. WHO. Available at: <https://www.who.int/tools/child-growth-standards/standards/weight-for-age>. Accessed 08.07.2023. (In Russ.)]
14. Клинические рекомендации. Ожирение у детей и подростков. Российская ассоциация эндокринологов. Доступно по: https://www.endocrinetr.ru/sites/default/files/specialists/science/clinic-recomendations/kr229_ozhirenie_u_detey_i_podrostkov.pdf. Ссылка активна на 08.07.2023. [Klinicheskie rekomendacii. Ozhirenie u detej i podrostkov. Rossijskaja asociacija jendokrinologov. Available at: https://www.endocrinetr.ru/sites/default/files/specialists/science/clinic-recomendations/kr229_ozhirenie_u_detey_i_podrostkov.pdf. Accessed 08.07.2023. (In Russ.)]
15. Порецкова Г.Ю., Тяжева А.А., Рапопорт И.К., Воронина Е.Н. Современные тренды нарушения здоровья детей школьного возраста г. Самара // Наука и инновации в медицине. — 2019. — №4(1). — С.58-62. [Poreckova GJu, Tjazheva AA, Rapoport IK, Voronina EN. Modern trends in the violation of the health of school-age children in Samara. Nauka i innovacii v medicine. 2019; 4(1): 58-62. (In Russ.)]
16. Соколов Д.К., Слазнева Т.И., Индершиев В.А., Абдуллаева Д.С. Ожирение как фактор риска повышения артериального давления у подростков // Молодой ученый. — 2020. — №23(313). — С.80-85. [Sokolov DK, Slazhneva TI, Indershiev VA, Abdullaeva DS. Ozhirenie kak faktor riska povysheniya arterial'nogo davleniya u podrostkov. Molodoj uchenyj. 2020; 23(313): 80-85. (In Russ.)]
17. Жилыев П.С., Горюнова Т.И., Завьялова Д.А. Внедрение информационно-аналитической системы «БАРС. WEB-МОНИТОРИНГ ЗДРАВООХРАНЕНИЯ» // Материалы VI Международной студенческой научной конференции «Студенческий научный форум». [ZHilyaev PS, Goryunova TI, Zav'yalova DA. Vnedrenie informacionno-analiticheskoy sistemy «BARS. WEB-MONITORING ZDRAVOOHRANE-NIYA». Materialy VI Mezhdunarodnoj studencheskoj nauchnoj konferencii «Studencheskij nauchnyj forum». (In Russ.)]
18. Евсикова Е.Ю., Курбатов М.Г. Возможности ранней диагностики гинекологической патологии у школьников с помощью автоматизированной системы АСПОН // Известия ЮФО. Технические науки. Доступно по: <https://cyberleninka.ru/article/n/vozmozhnosti-ranney-diagnostiki-ginekologicheskoy-patologii-u-shkolnits-s-pomoschyu-avtomatizirovannoy-sistemy-aspon/viewer>. Ссылка активна на 16.09.2023. [Evsikova EJU, Kurbatov MG. Vozmozhnosti ranney diagno-stiki ginekologicheskoy patologii u shkol'nic s pomoshch'yu avtomatizirovannoj sistemy ASPON. Izvestiya YUFO. Tekhnicheskie nauki. Available at: <https://cyberleninka.ru/article/n/vozmozhnosti-ranney-diagnostiki-ginekologicheskoy-patologii-u-shkolnits-s-pomoschyu-avtomatizirovannoy-sistemy-aspon/viewer>. Accessed 16.09.2023. (In Russ.)]
19. Шилков В.И., Самарцев А.А., Жовнер И.В. и др. Автоматизированная система профилактических осмотров (АСПОН-Д) как технология регистрации изменений показателей здоровья детей // Вестник уральской медицинской академической науки. — 2010. — №3. — С.94-96. [SHilkov VI, Samaracev AA, Zhovner IV, et al. Avtomatizirovannaya sistema profilakticheskikh osmotrov (ASPON-D) kak tekhnologiya registracii izmenenij pokazatelej zdorov'ya detej. Vestnik ural'skoj medicinskoj akademicheskoy nauki. 2010; 3: 94-96. (In Russ.)]

