

ГЕМАТОЛОГИЯ



© КОЛЛЕКТИВ АВТОРОВ, 2025

Вареха Н.В.¹, Стуклов Н.И.¹, Гимадиев Р.Р.^{1,2,3}, Гордиенко К.В.³, Щеголев О.Б.²,
Макарчев А.И.³, Гуркина А.А.¹

<https://elibrary.ru/rfxkfc>

СОЗДАНИЕ И ОЦЕНКА ЗНАЧИМОСТИ ПРОГНОСТИЧЕСКИХ МОДЕЛЕЙ ДЛЯ ОПРЕДЕЛЕНИЯ УРОВНЯ ФЕРРИТИНА СЫВОРОТКИ С ПОМОЩЬЮ МАШИННОГО ОБУЧЕНИЯ В РАЗНЫХ КЛИНИЧЕСКИХ ГРУППАХ

¹ФГАОУ ВО «Российский университет дружбы народов имени Патриса Лумумбы», 117198, Москва, Россия;

²ООО «ЛабХаб», 119002, Москва, Россия;

³ГНЦ РФ - ИМБП РАН, 123007, Москва, Россия

Изучение возможностей искусственного интеллекта (ИИ) является актуальной задачей с точки зрения разработки оптимальной скрининг стратегии, определения групп риска, использования менее дорогостоящих, более доступных лабораторных тестов для оценки статуса железа в организме.

Цель работы: изучить эффективность применения инструментов машинного обучения (МО) для оценки статуса железа по прогнозируемому уровню ферритина сыворотки (ФС) на основе демографических данных (пол и возраст), клинического анализа крови (КАК), содержания С-реактивного белка (СРБ) и известного уровня ФС. Для проведения МО с помощью ИИ использован набор данных 52 158 пациентов. Полученные данные представлены в виде Первой модели регрессии для определения предполагаемой концентрации ФС и Второй модели классификации групп пациентов в зависимости от различного статуса железа по уровню известного ФС: 0) <15,0; 1) 15,1-100,0; 2) 100,1-300,0; 3) >300,1 мкг/л. В результате Первая модель получила адекватную предсказательную способность ($R^2=0,717$), при этом ее качество тем лучше, чем ниже значение ФС (средняя абсолютная ошибка составила 2,4 мкг/л для класса пациентов с ФС <15,0 мкг/л) на тестовой выборке. Вторая модель показала ещё более высокую диагностическую возможность с точностью для разных клинических групп (показатель AUC-ROC: 0,914, 0,807, 0,812, 0,891 соответственно), что важно для определения тактики ведения пациентов. В результате проведённого исследования можно сделать заключение, что расчёт содержания ФС с помощью разработанных в работе моделей может использоваться в качестве точного и клинически значимого инструмента для оценки статуса железа в реальной врачебной практике.

Ключевые слова: дефицит железа; железодефицитная анемия; искусственный интеллект; машинное обучение; сывороточное железо; ферритин

Для цитирования: Вареха Н.В., Стуклов Н.И., Гимадиев Р.Р., Гордиенко К.В., Щеголев О.Б., Макарчев А.И., Гуркина А.А. Создание и оценка значимости прогностических моделей для определения уровня ферритина сыворотки с помощью машинного обучения в разных клинических группах. *Клиническая лабораторная диагностика*. 2025; 70 (3): 172-181.
DOI: <https://doi.org/10.51620/0869-2084-2025-70-3-172-181>
EDN: RFXKFC

Для корреспонденции: Вареха Николай Вячеславович, аспирант каф. госпитальной терапии с курсами эндокринологии, гематологии и клинической лабораторной диагностики Медицинского института ФГАОУ ВО «Российский университет дружбы народов имени Патриса Лумумбы»; e-mail: niki2187@mail.ru

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Финансирование. Исследование не имело спонсорской поддержки.

Поступила 15.09.2024
Принята к печати 18.12.2024
Опубликовано 25.02.2025

Varekha N.V.¹, Stuklov N.I.¹, Gimadiev R.R.^{1,2,3}, Gordienko K.V.³, Shchegolev O.B.², Makarchev A.I.³, Gurkina A.A.¹

DEVELOPMENT AND VALUE EVALUATION OF PREDICTIVE MODELS FOR DETERMINING BLOOD SERUM FERRITIN LEVELS IN DIFFERENT CLINICAL GROUPS USING MACHINE LEARNING

¹Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University), 117198, Moscow, Russia;

²LabHub LLC, 119002, Moscow, Russia;

³Institute of Biomedical Problems of the Russian Academy of Sciences, 123007, Moscow, Russia

Studying the capabilities of artificial intelligence (AI) is important to develop an optimal screening strategy, identify risk groups, and create less expensive laboratory tests to assess the iron status. The aim of the study is to evaluate the effectiveness of using machine learning (ML) tools to assess the iron status by the predicted serum ferritin (SF) level based on demographic data (gender and age), complete blood count (CBC), C-reactive protein (CRP) content and the historic data on the SF level. To perform ML using AI, a dataset of 52,158 patients was used. The obtained data were presented in the form of the First regression model to determine the pre-

dicted SF concentration and the Second model for classifying patient groups depending on different iron status by the level of known SF: 0) <15.0, 1) 15.1-100.0, 2) 100.1-300.0, 3) >300.1 µg/L. As a result, the First model demonstrated adequate predictive ability ($R^2=0.717$), and its quality is better; the lower the SF value (the average absolute error was 2.4 µg/L for the class of patients with SF <15.0 µg/L) in the test sample. The Second model showed an even higher diagnostic ability with accuracy for different clinical groups (AUC-ROC indicator: 0.914, 0.807, 0.812, 0.891, respectively), which is important for determining patient management tactics. As a result of the study, it can be concluded that the determination of SF content using the models developed can be used as an accurate and clinically significant tool for assessing iron status in clinical practice.

Key words: iron deficiency; iron-deficiency anemia; artificial intelligence; machine learning; iron level; ferritin

For citation: Varekha N.V., Stuklov N.I., Gimadiev R.R., Gordienko K.V., Shchegolev O.B., Makarchev A.I., Gurkina A.A. Development and value evaluation of predictive models for determining blood serum ferritin levels in different clinical groups using machine learning.

Klinicheskaya Laboratornaya Diagnostika (Russian Clinical Laboratory Diagnostics). 2025; 70 (3): 172-181 (in Russ.).

DOI: https://doi.org/10.51620/0869-2084-2025-70-3-172-181

EDN: RFXKFC

For correspondence: Varekha Nikolai Vyacheslavovich, Ph.D. student, Department of Therapy with the courses of endocrinology, hematology, and clinical laboratory diagnostic of the Medical Institute, Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University); e-mail: niki2187@mail.ru

Information about authors:

Varekha N.V., <https://orcid.org/0009-0004-0969-6144>;

Stuklov N.I., <https://orcid.org/0000-0002-4546-1578>;

Gimadiev R.R., <https://orcid.org/0000-0002-9567-3317>;

Gordienko K.V., <https://orcid.org/0000-0002-5381-1013>;

Shchegolev O.B., <https://orcid.org/0000-0002-3493-1415>;

Makarchev A.I., <https://orcid.org/0009-0008-7840-2540>;

Gurkina A.A., <https://orcid.org/0000-0003-4164-0058>.

Conflict of interests. The study had no sponsor support.

Acknowledgment. The authors declare absence of conflict of interests.

Received 15.09.2024

Accepted 18.12.2024

Published 25.02.2025

Введение. Дефицит железа (ДЖ) у детей младшего возраста и на ранних этапах беременности является актуальной проблемой [1]. Он увеличивает восприимчивость к инфекциям, повышает уровень материнской смертности и увеличивает длительность периода нетрудоспособности больных [2], экономические издержки, связанные с лечением и ведением этого синдрома, создают нагрузку на систему общественного здравоохранения и национальную экономику [3].

Для выявления ДЖ в медицинской практике чаще всего применяется анализ содержания ФС, считающегося наиболее объективным показателем. Анализ ФС плохо гармонизирован, при определении референсных интервалов и интерпретации значений концентрации необходимо учитывать различия между методами лабораторных исследований и анализаторами [4]. Диагностика ДЖ осложнена выраженным влиянием воспаления и других патологических состояний на показатели ФС у конкретного пациента.

Лабораторное определение ФС повсеместно не проводится при оказании первичной медико-санитарной помощи, в связи с этим нет возможности точной интерпретации ранних изменений в рутинных общеклинических и биохимических лабораторных исследованиях обмена железа, что особенно актуально для групп риска по ДЖ, в которые входят все женщины репродуктивного возраста и беременные. В последние годы исследователи уделяли особенное внимание возможностям МО и системы поддержки принятия врачебных решений (СППВР) для диагностики железодефицитных состояний. Модели МО могут обучаться на основе данных лабораторных исследований, запоминать информацию

и делать прогнозы. СППВР активно развиваются, особенно в области диагностики, поскольку они быстрые, точные, экономичные. Создаваемые модели превосходят существующие алгоритмы в точности оценки различных железодефицитных состояний [5-9].

Современное развитие ИИ даёт новые возможности МО для создания высокоточных прогностических моделей по определению ФС. Разработка таких моделей поможет быстро и точно, тем более в рамках рутинного первичного обследования, давать необходимую информацию по статусу железа, что может значимо улучшить оказание медицинской помощи населению.

Материал и методы. Для решения поставленной цели из лабораторной информационной системы сетевых лабораторий составлена выборка из 52 158 обезличенных лабораторных отчётов по клиническому анализу крови (КАК), С-реактивному белку (СРБ), ФС, сведения о поле и возрасте пациента (36 052 пациенток женского и 16 106 пациентов мужского пола).

Лабораторные исследования. КАК выполнен на автоматических гематологических анализаторах **УКАЗАТЬ в скобках страну!** Sysmex, Beckman Coulter, Mindray (КНР); скорость оседания эритроцитов (СОЭ) определялась на анализаторах Alifax, Diesse, Beckman Coulter; уровень ФС и СРБ – на иммунохимических и биохимических анализаторах Roche, Beckman Coulter, Abbott.

Алгоритм машинного обучения. Поставленная цель работы состояла из двух задач.

1. Создание Первой модели по прогнозированию уровня ФС на основе демографических данных (пол и возраст), данных КАК, СРБ и известного ФС (задача регрессии).

2. С использованием тех же показателей создание Второй модели по определению клинических групп ДЖ – попаданию расчётного показателя ФС в определённый диапазон значений (задача классификации): группа 0 (абсолютный ДЖ) – известный уровень ФС <15,0 мкг/л (5745 лиц женского и 453 мужского пола); группа 1 (возможный ДЖ, особенно при наличии воспаления) - ФС от 15,1 до 100,0 мкг/л (23 351 и 5884 человек); группа 2 (без ДЖ) - ФС от 100,1 до 300,0 мкг/л (5763 и 6893 человек); группа 3 (назначение препаратов железа не рекомендуется – возможный избыток железа) - ФС >300,1 мкг/л (1193 и 2876 человек) [10-12].

Предобработанные результаты лабораторных измерений до построения моделей разделены на обучающую и тестовую выборки в соотношении 4:1 (41726 и 10432 пациентов).

Для решения задач регрессии уровня ФС и классификации клинических групп пациентов по уровню ФС использован фреймворк LightAutoML [13] – специализированное программное обеспечение на языке Python, позволяющее в автоматическом режиме оптимизировать гиперпараметры (настройки модели, оказывающие влияние на то, как модель будет обучаться, какие признаки будут учитываться, и какие ограничения будут наложены на процесс обучения) используемых алгоритмов и произвести их блендинг (композицию нескольких алгоритмов с различными параметрами). В задаче регрессии использована линейная регрессия и градиентный бустинг на решающих деревьях (GBDT). GBDT - ансамбль решающих алгоритмических деревьев, построенных последовательно таким образом, что каждое следующее уменьшает ошибку предыдущего. В задаче классификации использованы два ал-

горитма: логистическая регрессия [14] и GBDT [15]. Реализация линейных моделей (линейной и логистической регрессий) взята из библиотеки Pytorch [pytorch.org] с использованием алгоритма линейной оптимизации L-BFGS [16]. Реализация градиентного бустинга на решающих деревьях взята из библиотеки LightGBM [lightgbm.readthedocs.io]. Обучение производилось с использованием кросс-валидации на 5 частях (5-fold cross-validation) [17], то есть обучающий набор данных случайным образом разбит на 5 равных частей для более верной оценки качества моделей и выбора оптимальных гиперпараметров (рис. 1). В качестве метрик качества моделей использован коэффициент детерминации (R^2) и площадь под ROC-кривой (AUC ROC) для задачи регрессии и классификации, соответственно.

Статистическая обработка. Для анализа различий между группами использованы следующие статистические методы. Для проверки гомогенности дисперсий применён тест Брауна-Форсайта (Brown-Forsythe). В случае выявления неоднородности дисперсий использован метод анализа дисперсий Уэлша (Welsh ANOVA, однофакторный дисперсионный анализ). Для последующих множественных сравнений применена поправка Геймса-Хауэлла (Games-Howell). Помимо статистической значимости различий между группами оценена практическая значимость (размер эффекта) с помощью коэффициента Коэна (d). Корреляция между исследуемыми параметрами определена с помощью критерия Пирсона. Различия считали статистически значимыми при $p < 0,05$ и практически значимыми при $d \geq 0,5$ [18]. Сравнение групп выполнено с помощью программного обеспечения GraphPad Prism 8, коэффициент Коэна рассчитан отдельно.

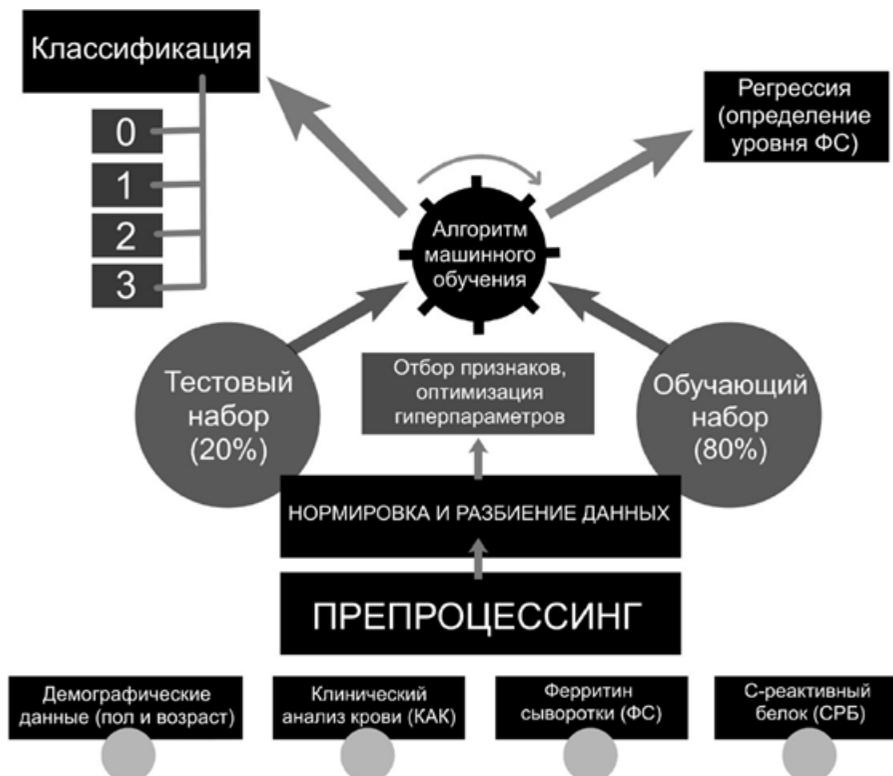


Рис. 1. Дизайн исследования, где классы пациентов распределены по уровню известного ФС: 0) <15,0; 1) 15,1-100,0; 2) 100,1-300,0; 3) >300,1 мкг/л.

Результаты. С целью определения наиболее значимых для разрабатываемых прогностических моделей параметров на первом этапе проведена оценка и взаимосвязь полученных лабораторных показателей и

демографических данных в общей когорте пациентов (Σ) и выборках с различными диапазонами ФС.

У лиц женского пола выявлена умеренная корреляционная связь между уровнем ФС и СРБ ($r=0,36$), рис. 2.

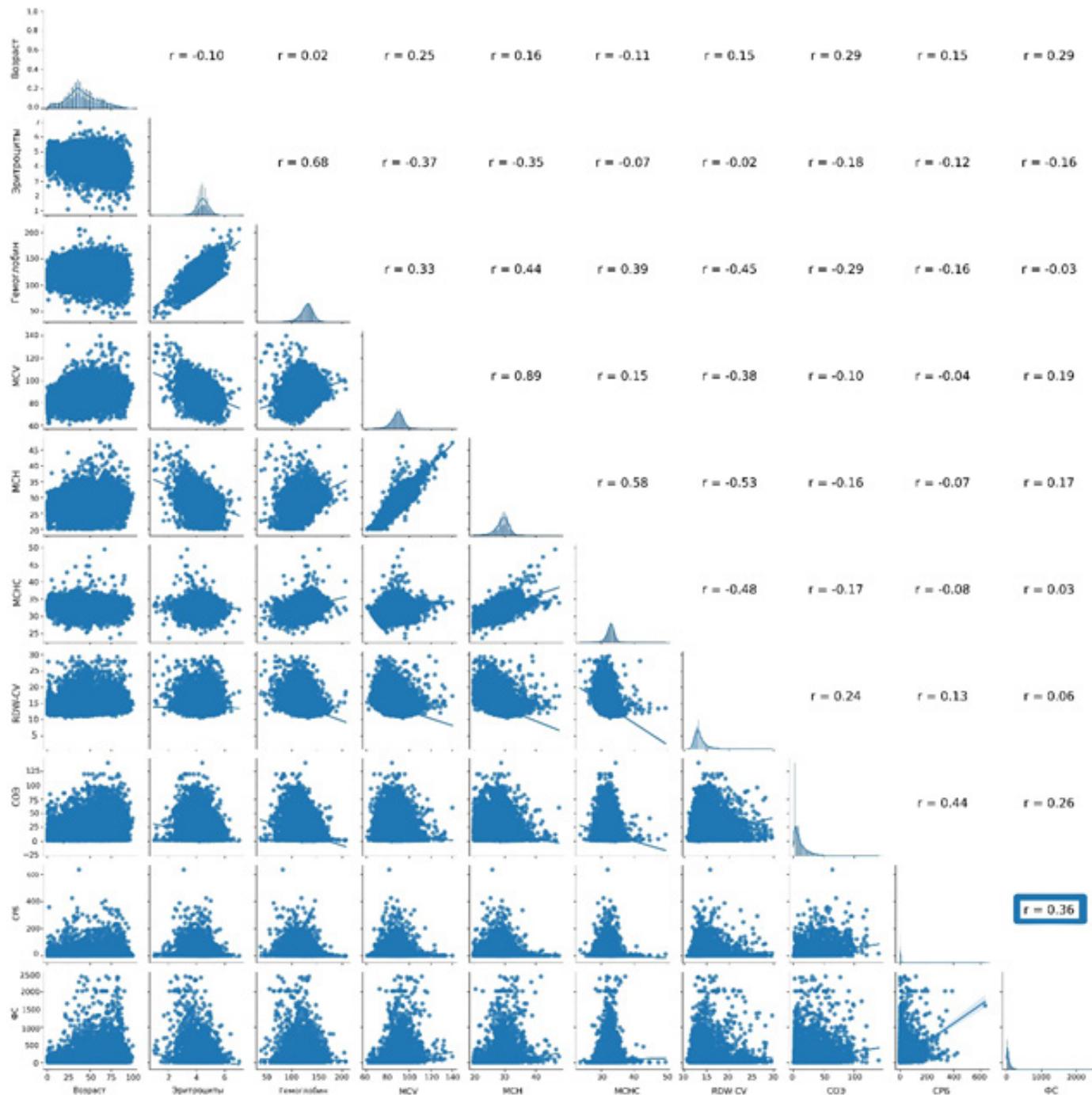


Рис. 2. Парный график исследуемых показателей у лиц женского пола.

У лиц мужского пола выявлена умеренная корреляционная связь между известным уровнем ФС и СРБ ($r=0,33$), ФС и возрастом ($r=0,30$) (рис. 3).

Как в общей когорте пациентов Σ , так и в выборках с различными диапазонами известных значений ФС выявлены статистически значимые отличия по-

казателей КАК, СРБ, ФС в зависимости от пола пациента ($p<0,05$; табл. 1), однако эффект практически незначим ($d<0,5$): исключение составило количество эритроцитов, половые различия по которому достигали величины в 1,4 стандартных отклонения (SD).

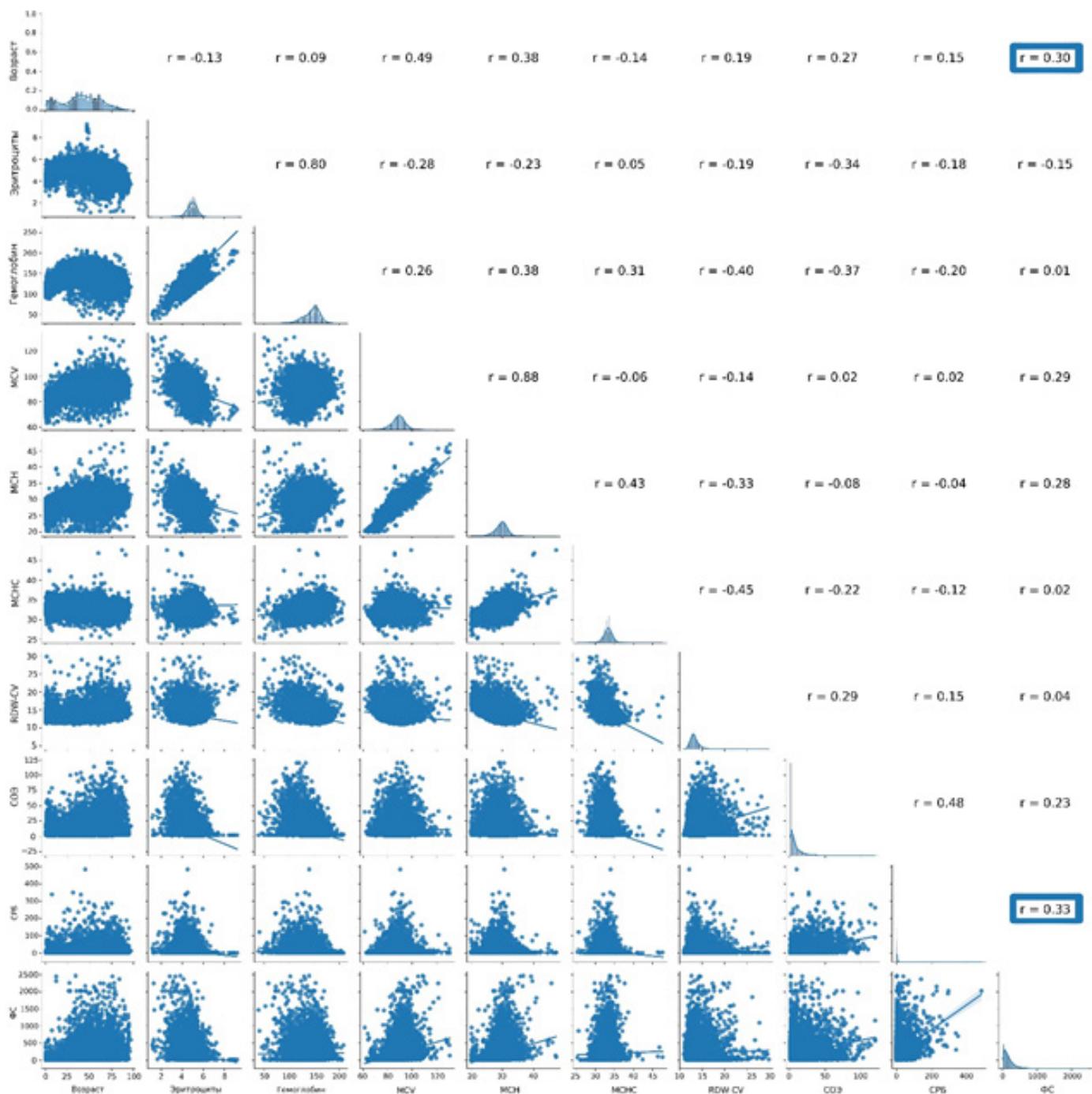


Рис. 3. Парный график исследуемых показателей у лиц мужского пола.

В разрезе диапазонов по уровню известных значений ФС выделенные интервалы демонстрировали прямо пропорциональную связь с возрастом: меньшей концентрации соответствовал меньший возраст. Практически значимым являлось отличие первой пары групп (пара групп 0 и 1, $ФС \leq 100$) от второй (пара групп 2 и 3, $ФС > 100,1$). Подобная картина сохранялась вне зависимости от пола. Возраст женщин в среднем был больше возраста мужчин, однако величина эффекта была средней ($d < 0,8$).

Высокая вариабельность СРБ и СОЭ существенно

снизила практическую значимость различий между исследуемыми группами по указанным показателям. Для женщин величина показателя позволяет выделить группу 3 с возможным избытком железа. Для мужчин подобная разница менее очевидна.

Уровень гемоглобина достоверно ниже у женщин по сравнению с мужчинами в исследованной выборке (что и должно быть в здоровой популяции), при этом эффект различия возрастает в зависимости от известного уровня ФС (прямо пропорционально).

Внутри половые различия по измеренному лабо-

раторно ФС между группами, однонаправленные для мужчин и женщин, следующие: при самом низком значении в группе 0, увеличении для 1-й и затем 2-й группы, наблюдается снижение показателя для группы 3, более выраженное для женщин. Концентрация ФС ожидаемо различалась между группами, поскольку выделение групп производилось по этому показателю. В рамках каждой группы значение показателя достоверно ниже для женщин, но эффект незначительный.

Для среднего содержания гемоглобина в эритроците (показателя МСН) статистически и практически значимым было отличие группы 0 от других групп вне зависимости от пола. В то же время для мужчин средний эффект обнаружен для различий между группой 1 и группами 2 и 3. Подобная ситуация наблюдалась и для средней концентрации гемоглобина в эритроцитах (МСНС), но значения эффектов менее выражены или отсутствуют вовсе.

В исследуемой когорте средний объём эритроцитов (MCV) у пациентов женского пола в среднем выше, чем у мужского, вне зависимости от группы по ФС, но со снижением размера эффекта при большей его концентрации. Как для мужчин, так и для женщин груп-

па 0 наиболее выражено отличается от других групп, при этом обособленность группы 1 более выражена для мужчин, нежели для женщин.

Абсолютное содержание эритроцитов вне зависимости от группы у пациентов женского пола ожидаемо отличалось от лиц мужского пола. При сравнении групп практически значимыми явились отличия группы 3 от других групп, но только для женщин.

Ширина распределения эритроцитов по объёму (RDW-CV) частично зависит от пола пациента: для женщин значения показателя для группы 0 и группы 3 более выражены, чем для групп 1 и 2; для мужчин группы 1, 2, 3 не имеют практически значимых отличий и значимо ниже по ширине распределения эритроцитов по сравнению с группой 0.

По исследованным характеристикам различных лабораторных и демографических показателей получены данные о значимой роли как КАК, так и содержания СРБ и известного ФС, их взаимосвязь в разных клинических и гендерных группах, что указывает на то, что все исследованные параметры должны в обязательном порядке быть включены в разработку прогностических моделей по определению содержания прогнозируемого ФС с помощью алгоритмов МО.

Таблица 1

Лабораторные показатели и демографические данные в общей когорте пациентов (Σ) и в выборках с различными диапазонами ФС в зависимости от пола

Показатель	Выборки	Женщины		Мужчины		Значимость
		M \pm SD	Me (Q ₁ ; Q ₃)	M \pm SD	Me (Q ₁ ; Q ₃)	
Ферритин, мкг/л	Σ	76,3 \pm 132,43	42,5 (21,3; 82,4)	194,8 \pm 222,27	132,5 (62,4; 244)	$p < 0,001$ $d = 0,05$
	0	9,7 \pm 3,21	10 (7,2; 12,4)	10,6 \pm 2,94	10,9 (8,5; 13)	$p < 0,001$ $d = 0,27$
	1	45,5 \pm 22,41	40,6 (26,8; 60,9)	54,4 \pm 24,55	52,2 (33; 75,8)	$p < 0,001$ $d = 0,39$
	2	158 \pm 49,8	143,2 (118; 187,2)	180,1 \pm 54,91	171 (132,9; 222,8)	$p < 0,001$ $d = 0,42$
	3	606,5 \pm 393,58	450 (352,4; 694)	546,5 \pm 313,55	435,5 (354; 603)	$p < 0,001$ $d = 0,18$
Возраст, годы	Σ	41,1 \pm 18,69	39 (29; 53)	40,6 \pm 22	41 (25; 57)	$p = 0,008$ $d = 0,01$
	0	35,3 \pm 13,81 ^{BB}	35,1 (26; 44)	23,5 \pm 24,64 ^{BB}	11 (4; 43)	$p < 0,001$ $d = 0,79$
	1	38,7 \pm 18,17 ^{ГД}	37 (28; 49)	31 \pm 24,69 ^{ГД}	25 (9; 53)	$p < 0,001$ $d = 0,39$
	2	52,8 \pm 17,89	55 (40; 65)	45,4 \pm 17,04	44 (33,1; 58)	$p < 0,001$ $d = 0,42$
	3	60,9 \pm 17,42	63 (49; 74)	51,3 \pm 16,41	49 (39; 62,3)	$p < 0,001$ $d = 0,57$
СРБ, мг/л	Σ	5,4 \pm 18,15	1,2 (0,4; 3,6)	6,5 \pm 20,16	1,3 (0,5; 3,6)	$p < 0,001$ $d = 0,01$
	0	2,1 \pm 3,97 ^B	0,8 (0,3; 2,2)	1,7 \pm 5,71	0,4 (0,2; 1,2)	$p = 0,875$ $d = 0,09$
	1	3,5 \pm 9,37 ^{ГД}	1,1 (0,4; 3)	3 \pm 8,31 ^{ГД}	0,8 (0,3; 2,3)	$p = 0,004$ $d = 0,05$
	2	10,4 \pm 26,81 ^{ГД}	2,5 (0,8; 7,6)	5,5 \pm 14,55	1,4 (0,6; 3,7)	$p < 0,001$ $d = 0,23$
	3	35 \pm 59,76	7,9 (2; 39,3)	16,6 \pm 38,61	2,6 (1; 11,7)	$p < 0,001$ $d = 0,40$

Показатель	Выборки	Женщины		Мужчины		Значимость
		M±SD	Me (Q ₁ ; Q ₃)	M±SD	Me (Q ₁ ; Q ₃)	
СОЭ, мм/час	Σ	12,5±12,73	8 (4; 17)	9,7±11,92	6 (2; 12)	p<0,001 d=0,06
	0	12,3±11,39 ^В	9 (4; 17)	8,7±10,29	5 (2; 11)	p<0,001 d=0,32
	1	10,8±10,7 ^Д	7 (3; 14)	7,9±9,38 ^А	5 (2; 9)	p<0,001 d=0,27
	2	16,3±15,21 ^Е	12 (6; 22)	9,3±10,77	6 (2; 11)	p<0,001 d=0,54
	3	28,2±23,15	22 (10; 40)	14,5±17,07	8 (3; 18)	p<0,001 d=0,72
Гемоглобин, г/л	Σ	129,1±13,23	130 (122; 138)	144,9±17,65	148 (135; 157)	p<0,001 d=0,25
	0	118,1±13,2 ^{АВ}	119 (110; 127)	124±17,09 ^{АВВ}	123 (113; 133)	p<0,001 d=0,44
	1	131,2±10,88 ^А	132 (125; 138)	140,5±16,78 ^Г	142 (129; 153)	p<0,001 d=0,75
	2	132,9±13,63 ^Е	134 (126; 142)	149,9±14,4	151 (143; 159)	p<0,001 d=1,21
	3	122,3±21,17	125 (109; 138)	145,2±21,4	150 (136; 159)	p<0,001 d=1,07
МСН, пг	Σ	29,2±2,35	29,4 (28; 30,6)	29,6±2,24	29,8 (28,4; 30,9)	p<0,001 d=0,11
	0	26,8±2,73 ^{АВВ}	27,2 (25,1; 28,7)	25,6±2,58 ^{АВВ}	26 (24,1; 27,2)	p<0,001 d=0,45
	1	29,5±1,93	29,6 (28,4; 30,7)	28,7±2,17 ^{ГД}	28,9 (27,4; 30,2)	p<0,001 d=0,36
	2	30±1,94	30 (28,9; 31,1)	30,1±1,75	30,2 (29,2; 31,2)	p<0,001 d=0,10
	3	30,1±2,78	30 (28,6; 31,4)	30,5±2,14	30,4 (29,3; 31,5)	p=0,003 d=0,15
МСНС, г/дл	Σ	32,8±1,23	32,9 (32,1; 33,6)	33,3±1,22	33,4 (32,6; 34,1)	p<0,001 d=0,37
	0	32±1,42 ^{АБ}	32,1 (31,2; 32,9)	32,3±1,55 ^{АВВ}	32,3 (31,4; 33,3)	p=0,004 d=0,20
	1	33±1,11	33,1 (32,4; 33,7)	33,3±1,17	33,3 (32,6; 34)	p<0,001 d=0,26
	2	32,9±1,14	33 (32,3; 33,6)	33,4±1,15	33,4 (32,8; 34,1)	p<0,001 d=0,43
	3	32,7±1,36	32,7 (31,9; 33,5)	33,4±1,31	33,4 (32,6; 34,2)	p<0,001 d=0,51
MCV, фл	Σ	88,8±5,97	89,2 (85,6; 92,5)	88,7±6,14	89 (85,3; 92,3)	p=0,080 d=0,01
	0	83,8±6,49 ^{АВВ}	84,2 (79,6; 88,3)	79,3±6,67 ^{АВВ}	79,4 (74,7; 83)	p<0,001 d=0,69
	1	89,3±5,14 ^А	89,6 (86,4; 92,5)	86,4±6,07 ^Г	86,7 (82,1; 90,5)	p<0,001 d=0,55
	2	91±5,34	91 (88; 93,9)	90,2±4,84 ^Д	90 (87,2; 93)	p<0,001 d=0,15
	3	92,1±7,54	91,4 (87,4; 95,5)	91,3±5,91	90,7 (87,8; 94,3)	p=0,024 d=0,12
Эритроциты, x10 ¹² /л	Σ	4,4±0,43	4,4 (4,2; 4,7)	4,9±0,56	4,9 (4,6; 5,2)	p<0,001 d=1,4
	0	4,4±0,41 ^В	4,4 (4,2; 4,7)	4,8±0,61	4,8 (4,5; 5,1)	p<0,001 d=1,03
	1	4,4±0,39 ^Д	4,4 (4,2; 4,7)	4,9±0,52	4,9 (4,6; 5,2)	p<0,001 d=1,02
	2	4,4±0,47 ^Е	4,5 (4,2; 4,7)	5±0,5	5 (4,7; 5,3)	p<0,001 d=1,09
	3	4,1±0,73	4,2 (3,6; 4,6)	4,8±0,74	4,9 (4,5; 5,3)	p<0,001 d=0,96
RDW-CV, %	Σ	13,7±1,7	13,3 (12,7; 14,1)	13,5±1,49	13,2 (12,7; 13,9)	p<0,001 d=0,1
	0	14,9±1,97 ^{АБ}	14,5 (13,4; 15,9)	15,2±2,17 ^{АВВ}	14,7 (13,6; 16,4)	p=0,028 d=0,17
	1	13,4±1,43	13,1 (12,6; 13,8)	13,6±1,54	13,3 (12,8; 14)	p<0,001 d=0,16
	2	13,5±1,57 ^Д	13,2 (12,6; 13,9)	13,3±1,17	13,1 (12,6; 13,7)	p<0,001 d=0,16
	3	14,6±2,61 ^Е	13,8 (13; 15,4)	13,6±1,69	13,2 (12,6; 14)	p<0,001 d=0,52

Примечание. А, Б, В, Г, Д, Е - статистически значимые отличия (p<0,05) между парами групп 0-1, 0-2, 0-3, 1-2, 1-3, 2-3, соответственно, при d≥0,8; а, б, в, г, д, е - то же при d≥0,5; М – среднее, SD – стандартное отклонение, Me – медиана, Q₁ и Q₃ – первый и третий квартиль, соответственно.

Характеристики полученной регрессионной модели. Разрабатываемая Первая модель является моделью регрессии и предназначена для определения про-

гнозируемого количественного уровня ФС у конкретного пациента. Метрики точности для оценки качества модели в общей когорте представлены в табл. 2.

Таблица 2

Характеристики регрессионной модели

R ²	MAE, мкг/л	MAPE, %	MSE, (мкг/л) ²
0,717	37,6	49,6	8701,0

Примечание. R² - коэффициент детерминации; MAE - средняя абсолютная ошибка; MAPE - средняя абсолютная ошибка в процентах; MSE - среднеквадратичная ошибка.

Выявлена сильная взаимосвязь между известным и прогнозируемым уровнем ФС ($r=0,888$, $p<0,05$). Модель обладает умеренной предсказательной способностью. Показатель оценки эффективности R² является удовлетворительным ($>0,5$). MAE, как среднее абсолютных разностей между измеренным и предсказанными значениями, составила 37,6 мкг/л. Ошибка (MAPE) составила 49,6% от фактического (измеренного) значения ФС. MSE отражает влияние больших ошибок на качество модели и составляет 8701,0 (мкг/л)² RMSE (корень из среднеквадратичной ошиб-

ки в абсолютном значении) составляет 93,3 мкг/л. Перечисленные показатели качества разработанной модели рассчитаны на всём диапазоне измерения известных концентраций ФС, где на этапе машинного обучения преобладали, в том числе и экстремально высокие значения.

Характеристики разработанной классификационной модели. Вторая классификационная модель, разрабатываемая для наглядной оценки средней абсолютной ошибки в разных клинических группах, показала следующие характеристики (табл. 3).

Таблица 3

Характеристики регрессионной модели в разрезе диапазонов уровней ФС

Клиническая группа	MAE, мкг/л	MAPE, %	MSE, (мкг/л) ²
0	2,4	30,1	11,7
1	16,5	41,3	451,8
2	41,9	24,0	2827,1
3	197,0	29,1	110333,2

Примечание. MAE - средняя абсолютная ошибка; MAPE - средняя абсолютная ошибка в процентах; MSE - среднеквадратичная ошибка.

В качестве основной метрики модели используется показатель площади под кривой AUC ROC-анализа [19] (рис. 4).

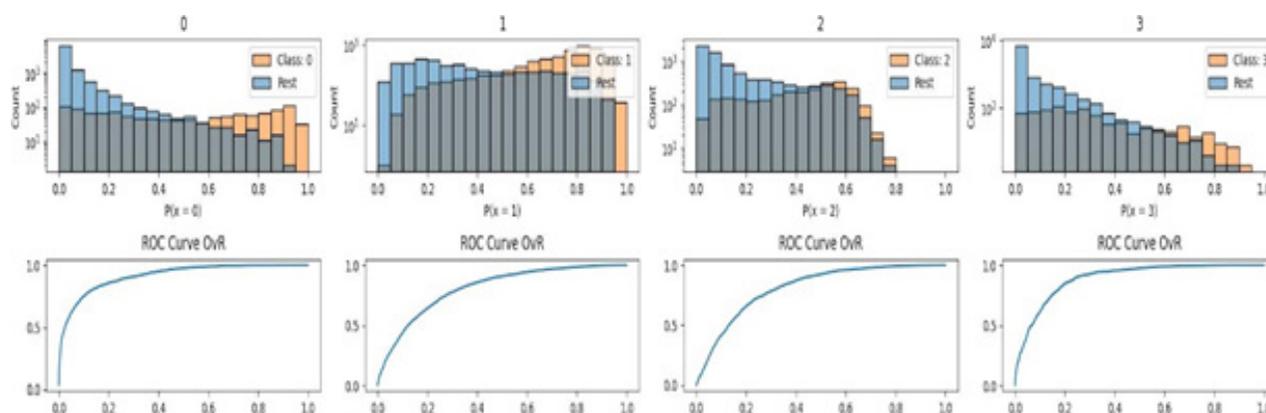


Рис. 4. ROC-кривые для 4 классов модели, рассчитанные по методу One-vs-Rest (OvR).

В табл. 4 представлена диагностическая точность классификационной модели в зависимости от пола пациентов.

Характеристики классификационной модели в разрезе пола пациентов

Параметры	Группа 0		Группа 1		Группа 2		Группа 3	
	Ж	М	Ж	М	Ж	М	Ж	М
ROC AUC (OvR)*	0,896	0,919	0,755	0,808	0,787	0,726	0,886	0,776
Чувствительность, %	52,5	26,7	92,5	56,6	16,0	82,7	26,2	22,1
Специфичность, %	96,5	99,2	37,9	87,8	96,7	52,1	99,2	95,9

Примечание. AUC ROC - площадь под ROC-кривой, рассчитанная по методу One-vs-Rest (OvR); М – мужчины; Ж - женщины; * - $p < 0,05$.

Показана достаточно высокая точность прогнозирования уровня ФС в разных клинических группах: AUC ROC составила 0,755-0,896 для лиц женского пола, 0,726-0,919 – для лиц мужского пола ($AUC \geq 0,7$) [20].

Обсуждение. ДЖ является одним из самых распространенных патологических состояний в мире, но до сих пор остаётся часто недиагностированным. Это связано как с нехваткой средств на проведение детального исследования статуса железа, так и с отсутствием возможности повсеместного использования точных лабораторных методов и их интерпретации, большой зависимостью полученных показателей от клинического статуса конкретного пациента (тем более, что само определение ФС не стандартизовано, и в лабораториях используют разные референсные интервалы). В это же время, когда медицинские данные в электронной карте пациента становятся все более структурированными, расширяются возможности МО и инструментов СППВР, особенно в области лабораторной медицины, поскольку они быстрые, точные и дешевые. Всё большее значение приобретают новые прогностические модели по определению статуса железа. Существует много алгоритмов предсказания содержания ФС с использованием как одного лабораторного параметра, так и множества (до 20), как лабораторных, так и демографических и клинических показателей.

При анализе лабораторных и демографических данных в общей выборке и подгруппах с различными диапазонами измеренного ФС стала понятна связь данного показателя с уровнем СРБ и возрастом (для мужского пола), что подтвердило необходимость обязательного включения указанных критериев как в регрессионную, так и тем более, в классификационную прогностическую модель по определению статуса железа. В результате первого этапа исследования получены данные, что все параметры статистически значимо отличаются между пациентами женского и мужского пола, что свидетельствует о разном уровне гендерной значимости проблемы и о необходимости дифференцированной разработки описываемых прогностических моделей.

Модель регрессии на тестовой выборке показала обнадеживающие результаты: $R^2=0,717$, $MAE=2,4$ мкг/л (для пациентов с ФС $<15,0$ мкг/л). Выявлена сильная корреляция между измеренным и прогнозируемым уровнем ФС ($r=0,888$), что согласуется с данными Y. Luo и соавт. ($r=0,732$) [6].

Результаты классификационной модели также показали высокую диагностическую точность (показатель ROC AUC от 0,726 до 0,919), что важно для определения тактики ведения больных. В связи с этим данная модель представляет существенный клинический интерес и может быть внедрена в медицинскую лабораторию.

Так, например, по результатам схожего исследования в нескольких регионах Нидерландов появилась услуга «рефлексного» тестирования. Специалисты лабораторной медицины могут вносить дополнительные тесты (ФС по результатам КАК и СРБ) в список необходимых для той или иной пробы по результатам предсказания алгоритма. Это экономит время врачей и сокращает количество повторных венопункций [8].

В нашем исследовании представлены результаты разработки двух прогностических моделей с высокой достоверностью предсказания концентрации ФС, полученных с помощью МО по анализу демографических данных, результатов КАК, СРБ и известных показателей ФС. Алгоритм, разработанный в этом исследовании, может быть использован как часть СППВР для выявления случаев дефицита железа в реальной клинической практике.

ЛИТЕРАТУРА (П.П. 1, 3 - 18 СМ. REFERENCES)

- Полякова О.А., Клепикова М.В., Литвинова С.Н., Исаакян Ю.А., Остроумова О.Д. Проблема дефицита железа и железодефицитной анемии в общей медицинской практике. *Профилактическая медицина*. 2022; 25(12):127-34.

REFERENCES

- <https://www.who.int/health-topics/anaemia>.
- Polyakova O.A., Klepikova M.V., Litvinova S.N., Isaakyan Yu.A., Ostroumova O.D. Iron deficiency anemia in general medical practice. *Profilakticheskaya meditsina*. 2022; 25(12):127-34. (in Russian)
- Blank P., Tomonaga Y., D. Szucs T., Schwenkglens M. Economic burden of symptomatic iron deficiency – a survey among Swiss women. *BMC Womens Health*. 2019; 19: 39.
- Hoofnagle A.N. Harmonization of blood-based indicators of iron status: making the hard work matter. *Am. J. Clin. Nutr.* 2017; 106:1615S-1619S.
- Azarkhish I., Raoufy M.R., Gharibzadeh S. Artificial intelligence models for predicting iron deficiency anemia and iron serum level based on accessible laboratory data. *J. Med. Syst.* 2012; 36(3):2057-61.
- Luo Y., Szolovits P., Dighe A.S., Baron J.M. Using machine learning to predict laboratory test results. *Am. J. Clin. Pathol.* 2016; 145(6):778-88.
- Pullakhandam S., McRoy S. Classification and explanation of iron deficiency anemia from complete blood count data using machine learning. *BioMedInformatics*. 2024; 4(1):661-72.
- Kurstjens S., de Bel T., van der Horst A., Kusters R., Krabbe J., van Balveren J. Automated prediction of low ferritin concentrations using a machine learning algorithm. *Clin. Chem. Lab. Med.* 2022; 60(12):1921-8.
- McDermott M., Dighe A.S., Szolovits P., Luo Y., Baron J.M. Using machine learning to develop smart reflex testing protocols. *J. Am. Med. Inform. Assoc.* 2023; 31(2):416-25.

10. WHO guideline on use of ferritin concentrations to assess iron status in individuals and populations. Geneva: World Health Organization; 2020.
11. Schop A., Stouten K., van Houten R., Riedl J., van Rosmalen J., Bindels P.J. et al. Diagnostics in anaemia of chronic disease in general practice: a real-world retrospective cohort study. *BJGP Open*. 2018; 2 (3):bjgpopen18X101597.
12. Weiss G., Goodnough L. Anemia of chronic disease. *New England Journal of Medicine*. 2005; 352(10):1011-23.
13. Vakhrushev A., Ryzhkov A., Savchenko M., Simakov D., Daminov R., Tuzhilin A. Lightautoml: Automl solution for a large financial services ecosystem. *arXiv preprint arXiv*. 2021; 2109.01528.
14. Nick T.G., Campbell K.M. Logistic regression. *Methods Mol. Biol*. 2007; 404:273-301.
15. Ke G., Meng Q., Finley T., Wang T., Chen W., Ma W. et al. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*. 2017; 30: 3149-57.
16. Moritz P., Nishihara R., Jordan M. A linearly-convergent stochastic L-BFGS algorithm. *Artificial Intelligence and Statistics*. 2016; 249-58.
17. Fushiki T. Estimation of prediction error by using K-fold cross-validation. *Statistics and Computing*. 2011; 21:137-46.
18. Diener M.J. Cohen's d. *The Corsini encyclopedia of psychology*. 2010; 1-1.
19. Yang Z., Xu Q., Bao S., Cao X., Huang Q. Learning with multiclass AUC: Theory and algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2021; 44 (11): 7747-63.
20. Swets J. A., Dawes R. M., Monahan J. Better decisions through science. *Scientific American*. 2000; 283(4):82-7.