

*На правах рукописи*

**Вареха Николай Вячеславович**

**Прогнозирование уровня ферритина сыворотки в общей популяции и при различных  
железодефицитных синдромах с использованием алгоритмов машинного обучения**

3.1.18. Внутренние болезни

Автореферат  
диссертации на соискание ученой степени  
кандидата медицинских наук

Москва – 2026

Работа выполнена на кафедре госпитальной терапии с курсами эндокринологии, гематологии и клинической лабораторной диагностики медицинского института федерального государственного автономного образовательного учреждения высшего образования «Российский университет дружбы народов имени Патриса Лумумбы» Министерства науки и высшего образования Российской Федерации

**Научный руководитель:**

доктор медицинских наук, доцент

**Стуклов Николай Игоревич**

**Официальные оппоненты:**

Рукавицын Олег Анатольевич - доктор медицинских наук, профессор, начальник гематологического центра ГВГК им Н.Н. Бурденко, главный внештатный гематолог Минобороны России.

Сахин Валерий Тимофеевич – доктор медицинских наук, начальник отделения реанимации и интенсивной терапии кардиоцентра ФГБУ «Национальный медицинский исследовательский центр высоких медицинских технологий им. А.А. Вишневского» Минобороны России.

**Ведущая организация** – федеральное государственное автономное образовательное учреждение высшего образования «Новосибирский национальный исследовательский государственный университет».

Защита диссертации состоится «18» июня 2026 г. в \_\_\_ часов на заседании диссертационного совета ПДС 0300.004 при ФГАОУ ВО «Российский университет дружбы народов имени Патриса Лумумбы» (117292, г. Москва, ул. Вавилова, д. 61, Университетская клиническая больница имени В. В. Виноградова (филиал) ФГАОУ ВО "Российский университет дружбы народов имени Патриса Лумумбы").

С диссертацией можно ознакомиться в читальном зале УНИБЦ (Научная библиотека) ФГАОУ ВО «Российский университет дружбы народов имени Патриса Лумумбы» по адресу: 117198, г. Москва, ул. Миклухо-Маклая, д. 6.

Автореферат разослан «\_\_» \_\_\_\_\_ 2026 г.

Ученый секретарь

диссертационного совета ПДС 0300.004,

доктор медицинских наук, профессор

**Сафарова Айтен Фуад кызы**

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы.** Дефицит железа (ДЖ) является основной причиной анемии во всем мире [Camaschella, 2019]. Согласно имеющимся данным, железодефицитная анемия (ЖДА), как крайняя степень выраженности ДЖ, диагностирована приблизительно у 1,92 миллиарда человек в мире [GBD 2021 Anaemia Collaborators, 2023]. Своевременное выявление ДЖ позволяет заранее, до наступления анемии и развития большого количества клинических проявлений сидеропенического синдрома, проводить профилактику его развития [Клинические рекомендации «Железодефицитная анемия», 2021].

Помимо классической ЖДА, связанной с абсолютным ДЖ (АДЖ), также выделяют функциональный ДЖ. Такие нарушения кроветворения, протекающие с перераспределением железа на фоне воспаления, определяют, как анемия хронической болезни (АХБ), и как сочетание ЖДА и АХБ – комбинированный ДЖ [Сахин В.Т. и др., 2023]. Основным лабораторным методом, на котором основано разделение всех вариантов анемий, протекающих с АДЖ или функциональным ДЖ, является определение уровня ферритина сыворотки (ФС).

Все указанные анемии развиваются в условиях железодефицитного эритропоэза, поэтому являются гипо-нормохромными микро- нормоцитарными с низким сывороточным железом и коэффициентом насыщения трансферрина железом, что с одной стороны, позволяет объединить их в единую группу железодефицитных синдромов (ЖДС), с другой – разделить пациентов с данной патологией на разные группы в зависимости от показателей ФС, причин развития анемии и демографических показателей. Такое разделение необходимо для принятия решения о диагностическом поиске и патогенетическом лечении анемии с или без использования препаратов железа [Стуклов Н.И. и др., 2024].

Несмотря на сложный процесс гомеостаза железа, до сих пор единственным критерием, позволяющим достаточно точно диагностировать различные варианты ЖДС, остается показатель ФС [Fletcher A. et al., 2022]. Применение ФС все еще не является скрининговым методом, используется далеко не всегда, обычно уже при наличии анемии для ее дифференциальной диагностики, ограничивается разными пороговыми значениями, широким диапазоном референсных значений, разницей методов в существующих тест-системах, а также зависимостью показателей феррокинетики от степени и длительности воспаления, что требует в спорных случаях использовать дополнительные клинические ориентиры (например, возраст и пол, поиск основной причины анемии) и лабораторные показатели, такие как, например, скорость оседания эритроцитов (СОЭ) и С-реактивный белок (СРБ) [Сахин В.Т. и др., 2020].

Алгоритмы машинного обучения (МО) способны обучаться на основе входящих данных клинических и лабораторных исследований, хранить информацию, а также строить прогнозы. Доказано, что применение алгоритмов МО способно помочь врачам не пропустить, и с достаточной точностью оценить степень ДЖ, и спрогнозировать уровень ФС у пациентов [Mojumdar M.U. et al., 2024]. При этом предполагается, что такой подход действительно может существенно упростить и улучшить диагностическую точность, снизить рабочую нагрузку на медицинский персонал и сократить количество упущенных и/или не диагностированных случаев ДЖ с одной стороны, и помочь рутинно использовать разработанные модели для точной дифференциальной диагностики ЖДС с другой.

**Степень разработанности темы исследования.** Современные исследования последнего десятилетия демонстрируют высокую эффективность МО в нескольких направлениях. С помощью нейросетей достигнута высокая точность ( $R^2 = 0,93$ ) при прогнозировании сывороточного железа [Azarkhish I. et al., 2012], разработана модель для прогноза ФС с корреляцией  $r = 0,73$  между предсказанными и измеренными значениями [Luo Y. et al., 2016]. Многочисленные исследования подтверждают высокую точность МО при дифференциальной диагностике ЖДА,  $\beta$ -талассемии и АХБ: площадь под кривой ROC-анализа (AUC ROC) составила до 0,99, точность до 95,2% [Bellinger C. et al., 2015; Ayyıldız H. et al., 2020; Yıldız T.K. et al., 2021; Pullakhandam S. et al., 2024]. Создан алгоритм для автоматического назначения теста на ФС по данным клинического анализа крови (КАК) и СРБ (AUC ROC = 0,90), внедренный в Нидерландах [Kurstjens S. et al., 2022]. Доказано превосходство МО над стандартными протоколами до назначения исследований [McDermott M. et al., 2023]. Несмотря на достигнутые успехи, остаются нерешенные проблемы: отсутствие единого подхода, недостаточное использование в моделях МО маркеров воспаления, фокусировка большинства исследований исключительно на определении ДЖ без учета возможного его избытка при различных вариантах ЖДС. На момент планирования данного исследования в доступной литературе не найдено работ, посвященных созданию подобных прогностических алгоритмов на репрезентативной популяционной выборке в Российской Федерации, что обосновывает актуальность и необходимость проведения настоящего исследования.

**Цель исследования** – разработка и оценка возможности применения алгоритмов МО для определения статуса железа по прогнозируемому уровню ФС на основе демографических данных (пол и возраст), параметров КАК, концентрации СРБ и известного уровня ФС в общей выборке пациентов и у больных с различными вариантами ЖДС.

**Задачи исследования:**

1. Определение распространенности дефицита и избытка железа (ИЖ), определение взаимосвязи известного содержания ФС и демографических данных, показателей КАК, СРБ в общей выборке пациентов.
2. Определение распространенности различных вариантов ЖДС, поиск взаимосвязи известного содержания ФС и демографических, лабораторных данных в когорте пациентов с микро- нормоцитарными анемиями.
3. На основании демографических данных, показателей КАК, СРБ и известного содержания ФС разработать и оценить диагностическую точность полученных алгоритмов МО, обученных и протестированных на общей выборке пациентов (алгоритм регрессии) и в разных по полу, возрасту и известному содержанию ФС группах (алгоритм классификации).
4. На основании указанных выше параметров разработать и оценить диагностическую точность полученных алгоритмов МО, обученных и протестированных на когорте пациентов с микро- нормоцитарными анемиями (алгоритм регрессии) и в разных по полу, возрасту и известным вариантам ЖДС группах (алгоритм классификации).
5. Сопоставить точность полученных данных в общей выборке и в выборке с микро-нормоцитарными анемиями, а также в разных группах по известному уровню ФС в этих когортах.
6. Оценить диагностическую точность полученных алгоритмов МО на когорте пациентов с клиническим диагнозом ЖДА, выявить факторы, ограничивающие применение данной методики.

**Научная новизна работы.** Предлагаемое исследование сосредоточено вокруг важной задачи внедрения инструментов МО, как раздела искусственного интеллекта, для оценки возможности получения данных о статусе железа с помощью предсказания уровня ФС. Точная оценка показателей обмена железа имеет важное значение, как для ранней диагностики ДЖ, еще до развития анемии и клинических симптомов ДЖ, так и для дифференциальной диагностики и подбора оптимального обследования и лечения таких состояний, как ЖДА, ЖДА+АХБ и АХБ, а также для заболеваний, связанных с накоплением железа. Теоретическая значимость этого исследования заключается в изучении пересечения между методами МО, лабораторной диагностикой и клинической медициной. Используя демографические данные, КАК, концентрацию СРБ и известные уровни ФС на большой выборке данных и в различных группах по степени ДЖ и вариантам ЖДС, проведено исследование, направленное на построение прогностических алгоритмов, которые могут не только оценивать уровни ФС, но и

улучшать понимание сложных взаимосвязей в многомерных клинических данных в диагностике нарушений обмена железа и развития различных вариантов анемий.

**Теоретическая и практическая значимость полученных результатов.** Разработанные алгоритмы ведут к улучшению технологий скрининга оценки статуса железа, тем самым предоставив врачам мощные инструменты для быстрого и точного выявления лиц с субклиническим и клинически значимым изменением уровня железа в крови. Используя известные данные, поставщики медицинских услуг смогут внедрять более эффективные и экономичные методы скрининга как в условиях оказания первичной медико-санитарной, так и специализированной помощи. Прогнозирование уровня ФС с помощью МО позволит своевременно выявлять группы риска, организовывать необходимые профилактические, проводить таргетированные диагностические и лечебные мероприятия, снижая бремя заболеваний, связанных с дисбалансом железа в организме. На основании проведенной работы разработано WEB-приложение (<https://med.labhub.online/smarthub>) для использования в реальной клинической практике.

**Методология и методы исследования.** Проведено ретроспективное когортное исследование, одобренное этическим комитетом РУДН. Анализ основан на деидентифицированных данных 62 148 пациентов за 2017–2024 гг., полученных из информационных систем четырех сетевых лабораторий. Включались пациенты  $\geq 18$  лет с одномоментным определением КАК, СРБ и ФС. Сформированы три выборки: общая ( $n=62\ 148$ ), пациенты с микро- нормоцитарными анемиями ( $n=9\ 771$ ) и внешняя валидационная ( $n=100$ ).

Формирование групп проведено по уровню ФС, в общей выборке выделены: абсолютный ДЖ (ФС  $<15$  мкг/л), возможный ДЖ (ФС 15–100 мкг/л), отсутствие ДЖ (ФС 100–300 мкг/л), ИЖ (ФС  $\geq 300$  мкг/л). В когорте с анемиями: ЖДА (ФС  $<15$  мкг/л), ЖДА+АХБ (ФС 15–100 мкг/л), АХБ (ФС 100–300 мкг/л), анемия без ДЖ (ФС  $\geq 300$  мкг/л). Все группы стратифицированы по полу и возрасту ( $<45$  и  $\geq 45$  лет). Модели протестированы на независимой выборке пациентов с верифицированным диагнозом ЖДА, в которой проведен анализ клинических случаев с максимальными расхождениями прогноза для выявления ограничений метода.

С помощью МО решали две задачи: регрессия (количественный прогноз ФС) и классификация (отнесение к диапазону ФС). Использован фреймворк LightAutoML с алгоритмами линейной регрессии и градиентного бустинга (LightGBM). Данные разделены на обучающую и тестовую выборки (4:1) с пятикратной кросс-валидацией.

Оценку качества полученных данных проводили методом регрессии, когда определяли коэффициент детерминации ( $R^2$ ), медианную абсолютную ошибку (MedAE), среднюю абсолютную ошибку в процентах (MAPE), корень из среднеквадратичной ошибки (RMSE) и методом классификации с получением данных AUC ROC, определением чувствительности и специфичности. Статистическую обработку проводили в GraphPadPrism 8.

#### **Положения, выносимые на защиту:**

1. ДЖ (ФС <100 мкг/л) выявляется в общей популяции чаще, чем ИЖ (61,2% и 12,0% соответственно,  $p < 0,001$ ), определена высокая частота ДЖ именно у женщин (75,8%), особенно в репродуктивном возрасте (89,5%), что делает эту группу пациентов наиболее подверженной риску развития анемии. Для точного прогнозирования уровня ФС в общей популяции, как и ожидалось, наиболее значимыми факторами определены возраст, лабораторные маркеры воспаления, гемоглобин (HGB) и параметры эритроцитов.
2. ЖДА и ЖДА+АХБ (ФС <100 мкг/л) у пациентов с микро- нормоцитарными анемиями имеют выраженную гендерно-возрастную специфичность и выявлены у подавляющего большинства лиц женского пола (80,6%), достигая максимума в репродуктивном возрасте (93,5%), в то время как у мужчин их распространенность существенно ниже (31,3%,  $p < 0,001$ ). При этом противопоказания к терапии препаратами железа (ФС  $\geq 300$  мкг/л) наиболее характерны для пациентов мужского пола (39,0%) и крайне редки у женщин, особенно моложе 45 лет (1,7%). Для точного прогнозирования уровня ФС на выборке пациентов с микро- нормоцитарными анемиями, наиболее значимыми факторами, помимо указанных в общей выборке, определены показатель анизоцитоза эритроцитов (RDW-CV), количество эритроцитов (RBC), средняя концентрация гемоглобина в эритроците (MCHC) (для всех  $p < 0,001$ ).
3. Регрессионный анализ на общей выборке пациентов подтвердил адекватную предсказательную способность модели ( $R^2 = 0,71$ ). Классификационный анализ выявил, что алгоритм обладает наивысшей диагностической значимостью для группы женщин с возможным ДЖ (AUC ROC = 0,75, чувствительность 92,0%). Наименее эффективно алгоритм классифицировал пациентов с ИЖ, что проявляется в низкой чувствительности (16,5%) при сохранении высокой специфичности (99,3%).
4. Регрессионный анализ на выборке пациентов с микро- нормоцитарными анемиями подтвердил высокую предсказательную способность ( $R^2 = 0,70$ ). Классификационный анализ показал, что алгоритм наиболее эффективен для диагностики ЖДА у женщин (ФС <15 мкг/л, AUC ROC = 0,91, чувствительность 82,3%). Минимальная эффективность

алгоритма отмечена для группы с АХБ ( $100 \leq \text{ФС} < 300$  мкг/л), где наблюдается низкая чувствительность (26,9%) при сохранении высокой специфичности (95%).

5. Наиболее точные результаты разработанные алгоритмы показывают при наличии анемии, низких показателях ФС у пациенток женского пола. При разработке регрессионных и классификационных алгоритмов МО в когорте пациентов с микро- нормоцитарными анемиями результаты оказались более точными, чем в общей группе у лиц мужского и женского пола как моложе 45, так старше 45 лет включительно (все  $p < 0,001$ ).
6. Алгоритм МО, обученный и протестированный на когорте пациентов с установленным клиническим диагнозом ЖДА, продемонстрировал удовлетворительную точность прогноза ФС при внешней валидации выборки пациентов. Основными ограничениями применения данного алгоритма являются терапия препаратами железа, гемотрансфузии и значительный временной интервал между анализами КАК, СРБ и определением прогностического ФС (2-3 недели). Применение алгоритма МО требует обязательного учета клинического контекста и рекомендовано только для первичного скрининга при предполагаемом клиническом диагнозе ЖДА.

**Внедрение результатов исследования в практику.** Результаты работы используются на кафедре госпитальной терапии с курсами эндокринологии, гематологии и клинической лабораторной диагностики Медицинского института РУДН. Алгоритмы МО, внедрены в практическую работу отделения медицинской профилактики №3 ГБУЗ «Консультативно-диагностическая поликлиника №121 Департамента здравоохранения города Москвы» (Акт о внедрении результатов диссертационной работы от 19.08.2025).

**Степень достоверности и апробация результатов.** Достоверность обеспечивается репрезентативным объемом выборки (62 148 пациентов), длительным периодом сбора данных (2017–2024 гг.), использованием сертифицированного лабораторного оборудования, современных алгоритмов МО с кросс-валидацией и внешней валидацией на независимой выборке. Апробация проведена на кафедре госпитальной терапии с курсами эндокринологии, гематологии и клинической лабораторной диагностики Медицинского института РУДН. Материалы исследования были представлены докладами и обсуждались на научно-практических конференциях различного уровня: XVII Межрегиональной научно-практической конференции «Весенняя сессия Московского городского научного общества терапевтов: терапевтические прорывы» 22 мая 2024 года; XXX Всероссийская юбилейная научно-практическая конференция «Национальные традиции клинической лабораторной диагностики», 18 марта 2025 года; Научно-практическая конференция «Госпитальная терапия: вчера, сегодня,

завтра», 23 апреля 2025 года; XXX Международная научно-практическая конференция «Пожилой больной. Качество жизни», 30 сентября 2025г.

**Публикации.** По результатам диссертации опубликовано 7 работ, из них SCOPUS/WoS – 4, RSCI/BAK – 2, перечень РУДН – 1.

**Соответствие диссертации паспорту научной специальности.** Диссертация соответствует паспорту научной специальности 3.1.18. внутренние болезни по направлениям: 1 – изучение этиологии, патогенеза и распространенности состояний и заболеваний (анализ эпидемиологии ДЖ и ИЖ); 2 – изучение патофизиологических проявлений с использованием лабораторных методов (исследование взаимосвязей показателей крови с уровнем ФС); 3 – совершенствование дифференциальной диагностики болезней внутренних органов (разработка алгоритмов диагностики ЖДА, АХБ и их сочетания); 6 – использование системы интеллектуального анализа данных, методов искусственного интеллекта в клинической практике (создание прогностических моделей и WEB-приложения).

#### **Личный вклад автора**

Автором самостоятельно выполнил анализ проблемы, определил цель и задачи исследования, спланировал план и дизайн проведения исследования, провел сбор клинического материала, создал базу данных, осуществил анализ полученных результатов.

**Объем и структура диссертации.** Диссертация изложена на 173 страницах машинописного текста, в своем составе имеет 36 таблиц и 52 рисунка, 3 приложения. По структуре имеет оглавление, список сокращений, введение, обзор литературы, материалы и методы исследования, результаты исследования, обсуждение полученных результатов, заключение, выводы, практические рекомендации, список литературы из 118 источников (44 отечественных и 74 зарубежных).

#### **ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ**

**Материалы и методы исследования.** Данная работа представляет собой ретроспективное когортное исследование. Аналитической основой исследования стал деидентифицированный массив данных, охватывающий демографические и лабораторные показатели 62148 пациентов за 2017-2024 гг., полученный с помощью автоматической выгрузки датасета из лабораторных информационных систем четырех сетевых лабораторий. Собрана информация о поле и возрасте пациентов, количестве лейкоцитов (WBC), RBC, HGB, гематокрита (HCT), MCV, MCH, MCHC, RDW-CV, тромбоцитов (PLT), относительном содержании нейтрофилов (NEU%), эозинофилов (EOS%), базофилов (BAS%), лимфоцитов (LYM%), моноцитов (MON%), СОЭ, уровне СРБ и известной концентрации ФС.

- 1) Общая выборка пациентов составила 62148 обследуемых (41754 лиц женского и 20394 лиц мужского пола).
- 2) Для исследования частоты и вариантов нарушений метаболизма железа при анемии отдельно исследовали данные с основным лабораторным критерием ЖДС – микро- нормоцитарной анемией. Выборка составила 9771 запись (7987 лиц женского и 1784 лиц мужского пола).
- 3) Для проведения внешней валидации и апробации метода была использована база данных с установленным клиническим диагнозом ЖДА из ГБУЗ «Консультативно-диагностической поликлиники №121 Департамента здравоохранения города Москвы», состоящая из 100 пациентов женского пола.

Критерии включения в исследование:

1. Пациенты мужского и женского пола в возрасте старше 18 лет включительно;
2. Наличие в лабораторной информационной системе полного набора данных для каждого пациента, при этом анализы КАК, СРБ и ФС были взяты у пациентов в один день (одномоментно).
3. Для формирования когорты пациентов с подозрением на ЖДА дополнительно применялись критерии микро- нормоцитарной анемии: уровень HGB менее 130 г/л для лиц мужского пола и менее 120 г/л для лиц женского пола, MCV менее 100 фл.
4. Для проведения внешней валидации в исследование включались пациенты женского пола с установленным клиническим диагнозом ЖДА из ГБУЗ «КДП №121 ДЗМ».

Критерии невключения в исследование: возраст менее 18 лет, отсутствие полного перечня показателей, необходимых для прогноза (неполные или ошибочные данные).

**Лабораторные исследования.** КАК был выполнен на автоматических гематологических анализаторах производителей Sysmex (США), Beckman Coulter (США), Mindray (КНР), СОЭ измерялась на анализаторах Alifax (Италия), Diesse (Италия), Beckman Coulter (США), ФС и СРБ – на иммунохимических и биохимических анализаторах Roche (Германия), Beckman Coulter (США), Abbott (США) с использованием соответствующих сертифицированных реактивов.

**Характеристика выборки и принципы разделения пациентов на группы в общей выборке пациентов.** Пациенты по уровню ФС были разделены в общей выборке обследуемых пациентов следующим образом:

Группа I: АДЖ: ФС <15 мкг/л (n = 7594).

Группа II: Возможный ДЖ, особенно при наличии воспаления:  $15 \leq \text{ФС} < 100$  мкг/л (n = 30339).

Группы I и II при необходимости объединяли в группу пациентов с ДЖ.

Группа III: Отсутствие ДЖ (ОДЖ):  $100 \leq \text{ФС} < 300$  мкг/л ( $n = 16734$ ).

Группа IV: ИЖ:  $\text{ФС} \geq 300$  мкг/л ( $n = 7481$ ).

Все указанные группы дополнительно были разделены на лиц мужского и женского пола, младше 45 и старше 45 лет включительно.

**Характеристика выборки и принципы разделения на варианты железодефицитных синдромов в когорте с микро- и нормоцитарными анемиями.** По признаку наличия микро-нормоцитарной анемии из общей выборки выделены пациенты с ЖДС. Аналогично принципу разделения в общей выборке пациенты разделены по содержанию ФС на разные группы:

Группа Ia: ЖДА:  $\text{ФС} < 15$  мкг/л ( $n = 3762$ ).

Группа IIa: ЖДА+АХБ:  $15 \leq \text{ФС} < 100$  мкг/л ( $n = 3231$ ).

Группа IIIa: АХБ:  $100 \leq \text{ФС} < 300$  мкг/л ( $n = 1539$ ).

Группа IVa: Анемия без ДЖ:  $\text{ФС} \geq 300$  мкг/л ( $n = 1239$ ).

Все указанные группы дополнительно также были разделены на лиц мужского и женского пола, младше и старше 45 лет включительно.

**Показатели выборки пациентов с установленным клиническим диагнозом железодефицитная анемия.** Для проведения внешней валидации и апробации метода в реальной практике в исследование была включена группа пациентов с установленным диагнозом ЖДА. Для исключения влияния других вариантов анемии, исследуемая выборка пациентов с ЖДА была представлена исключительно лицами женского пола. Для проведения сравнительного анализа данная группа была стратифицирована по возрастному признаку на две подгруппы: пациенты в возрасте младше 45 лет ( $n = 62$ ); пациенты в возрасте 45 лет и старше ( $n = 38$ ).

**Обучение и тестирование алгоритмов машинного обучения.** Создание и обучение модели МО состояло из нескольких алгоритмов и этапов. 1. В качестве первого алгоритма МО использовали задачу регрессии – прогнозирование абсолютного значения ФС на основе демографических данных, КАК, СРБ и известного содержания ФС на 1) общей выборке и 2) у пациентов с микро- и нормоцитарными анемиями. 2. В качестве второго алгоритма МО использовали задачу классификации – прогнозирование ФС в известном диапазоне по результатам тех же показателей. Для этого был создан новый целевой фактор, содержащий информацию о группах по содержанию в том или ином диапазоне ФС, в зависимости от пола и возрастного диапазона 1) в общей выборке пациентов и 2) с микро- и нормоцитарными анемиями.

Этапность исследования состояла в том, что предобработанные результаты лабораторных измерений до построения алгоритмов были разделены на обучающую и тестовую выборки в случайном порядке в соотношении 4:1 для пациентов 1) общей выборки (49718:12430 пациентов) и 2) с микро- и нормоцитарной анемией (7816:1955 пациентов). Для решения задач регрессии и классификации по уровню ФС был использован фреймворк LightAutoML [Vakhrushev A. et al., 2021], позволяющий в автоматическом режиме оптимизировать гиперпараметры используемых алгоритмов и произвести их блендинг. В задаче регрессии использовались линейная регрессия и градиентный бустинг. В задаче классификации были использованы два алгоритма: логистическая регрессия [Nick T. et al., 2007] и градиентный бустинг на решающих деревьях [Ke G. et al., 2017]. В случае многоклассовой классификации использовали методы "один против всех" (one-vs-all), когда строится отдельная модель для каждого класса и "один против одного" (one-vs-one), когда строится отдельная модель для каждой пары классов. Реализация линейных моделей (линейной и логистической регрессий) взята из библиотеки pytorch с использованием алгоритма линейной оптимизации L-BFGS [Moritz P. et al., 2015]. Реализация градиентного бустинга на решающих деревьях взята из библиотеки LightGBM. Обучение производилось с использованием кросс-валидации на 5 частях [Fushiki T., 2011; Kurstjens S. et al., 2022], то есть обучающий набор данных был случайным образом разбит на 5 равных частей для более верной оценки качества моделей и выбора оптимальных гиперпараметров. Входные данные проходили автоматический препроцессинг: удаление выбросов и некорректных значений, нормализацию. Выходную переменную в случае задачи регрессии логарифмировали для уменьшения влияния очень высоких значений уровня ФС на функцию потерь. Далее выбирали гиперпараметры обучения: функция потерь, основная валидационная метрика. Для задачи регрессии основной функцией потерь была выбрана MedAE, а основной метрикой –  $R^2$ . Для задач классификации функция потерь – кроссэнтропия, метрика – AUC ROC-анализа [Yang Z. et al., 2021]. Модели обучали на данных с использованием кроссвалидации на 5 частей: данные делятся случайным образом на 5 частей и модель обучается 5 раз на 4 разных частях, после чего её точность измеряли на оставшейся 5 части. Таким образом, в процессе перебора гиперпараметров (оптимизации) проводили максимально достоверную оценку точности модели. После выбора гиперпараметров и обучения на всём тренировочном наборе модель проверяли на отложенной тестовой выборке. В алгоритмах бустинга на решающих деревьях показатель p-value, как статистический параметр, не указывался и не упоминался ранее в работах по МО [Aigner E. et al., 2014].

**Статистическая обработка результатов.** Статистический анализ проводился с применением методов описательной статистики. Предварительно для каждого количественного признака проверяли гипотезу о его соответствии нормальному распределению с помощью тестов Шапиро-Уилка и Д'Агостино-Пирсона [Das K.R. et al., 2016]. Сравнение выделенных категориальными параметрами (пол, диапазоны ФС) групп по количественным признакам выполняли с помощью теста Манна-Уитни с поправкой Беньямини-Хохберга. Значимыми считали различия при скорректированном поправкой уровне  $p < 0,05$  и размере эффекта (дельта Клиффа), абсолютное значение которого было больше 0,33 [Meissel K. et al., 2024]. Для оценки статистической связи между количественными переменными использовали коэффициент корреляции Спирмена ( $r_s$ ) с тем же уровнем значимости и порогом практической значимости  $|r_s| > 0,3$ . Прогностическая точность модели по анализу AUC ROC считалась статистически значимой: слабо при значениях 0,60 - 0,70, умеренно при 0,70 - 0,80, хорошо при 0,80 - 0,90, отлично при  $> 0,90$  [Swets J.A. et al., 2000]; коэффициент детерминации  $R^2$  считался достоверным при значении  $> 0,20$  [Gupta A. et al., 2024]. Все статистические тесты выполняли с помощью программного обеспечения (ПО) GraphPadPrism 8, размер эффекта рассчитывали отдельно.

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

**Определение распространенности дефицита и избытка железа, определение взаимосвязи известного содержания ферритина сыворотки с демографическими данными, показателями клинического анализа крови, С-реактивного белка в общей выборке пациентов.** Число обследуемых, имеющих АДЖ составило 7594 (12,2%), возможный ДЖ – 30429 (49,0%), ОДЖ - 7481 (12,0%). Таким образом, ДЖ (АДЖ, возможный ДЖ) выявлен у подавляющего количества обследуемых в общей выборке – у 61,2%, ОДЖ – у 26,8%, ИЖ – у 12,0% соответственно.

Число обследованных лиц женского пола репродуктивного возраста (до 45 лет) в общей выборке пациентов составило 24002 (57,5% от общего количества обследованных женщин). Среди них заметно преобладание доли лиц с ДЖ, (89,5%). Эту группу представляют обследованные с уровнем ФС, соответствующему АДЖ ( $n = 5482$ , 22,8%) и возможному ДЖ ( $n = 16005$ , 66,7%). В данной группе АДЖ и возможный ДЖ встречается статистически значимо чаще по сравнению с лицами старше 45 лет (коэффициент  $\chi^2 = 1474$  и  $1371$ ,  $p < 0,001$ ). ИЖ в этих группах определен всего у 1,3% обследуемых. Группа лиц женского пола от 45 лет и старше составила 17752 человек, или 42,5% от общего количества обследуемых пациенток. В описываемой группе ДЖ выявлен у 10169 человек, что составило 57,3% этой выборки. ОДЖ и

ИЖ встречаются статистически значимо чаще, чем у молодых (32,4% и 10,3%, коэффициент  $\chi^2 = 3554$  и  $1725$ , соответственно, для всех  $p < 0,001$ ). Полученные данные свидетельствуют о том, что в популяции лиц у женского пола молодого возраста преобладает АДЖ и возможный ДЖ, а старшего – ОДЖ и ИЖ.

Число обследованных лиц мужского пола молодого возраста (до 45 лет) в общей выборке пациентов составило 10645 (52,2% от общего количества обследуемых мужчин). В этой группе доля с ДЖ составила 38,0% ( $n = 4044$ ), где АДЖ ( $n = 404$ , 3,8%) и возможный ДЖ ( $n = 3640$ , 34,2%). АДЖ и возможный ДЖ у молодых мужчин встречается статистически значимо чаще по сравнению с немолодыми (коэффициент  $\chi^2 = 72,92$  и  $370,9$ ,  $p < 0,001$ ). Группа лиц мужского пола 45 лет и старше составила 9749 человека (47,8% от общего количества обследуемых мужчин). В описываемой группе ДЖ выявлен у 2323 человек (23,8%), ОДЖ ( $n = 4190$ , 43,0%), ИЖ ( $n = 3236$ , 33,2%). ИЖ у мужчин 45 лет и старше встречается статистически значимо чаще, чем у молодых (коэффициент  $\chi^2 = 468,3$ ,  $p < 0,001$ ). В популяции лиц мужского пола молодого возраста преобладает АДЖ и возможный ДЖ, а старше – ИЖ.

Согласно критерию ( $|r_s| > 0,3$ ) определена следующая значимость показателей в прогнозировании уровня ФС в общей выборке пациентов: возраст ( $r_s = 0,42$ ), СРБ ( $r_s = 0,35$ ), НГВ ( $r_s = 0,32$ ), МСV ( $r_s = 0,34$ ), МСН ( $r_s = 0,35$ ) для лиц женского пола; возраст ( $r_s = 0,32$ ), СРБ ( $r_s = 0,42$ ), МСV ( $r_s = 0,30$ ), МСН ( $r_s = 0,30$ ) – для лиц мужского пола.

**Определение распространенности различных вариантов железодефицитных синдромов, поиск взаимосвязи известного содержания ферритина сыворотки и демографических, лабораторных данных в когорте пациентов с микро- нормоцитарными анемиями.** В выборке пациентов с микро- нормоцитарными анемиями число обследованных с ЖДА составило 3762 (38,5%), ЖДА+АХБ – 3231 (33,1%) в когорте пациентов с микро- нормоцитарными анемиями ( $n = 9771$ ). Число обследованных с анемией без ДЖ составило 1239 (12,7%). Таким образом, ЖДА и ЖДА+АХБ выявлены у подавляющего количества пациентов в когорте пациентов с микро- нормоцитарными анемиями – у 71,6%, АХБ – у 15,7%, анемия без ДЖ – у 12,7% пациентов.

Число лиц женского пола в группе ЖДА в когорте пациентов с микро- нормоцитарными анемиями составило 3606 (45,1%), в ЖДА+АХБ – 2828 (35,5%). Число лиц женского пола в группе анемии без ДЖ (ФС  $\geq 300$  мкг/л) составило 544 (6,8%). Таким образом, ЖДА и ЖДА+АХБ выявлены у подавляющего количества пациенток с микро- нормоцитарными анемиями – у 80,6%, АХБ – у 12,6%, анемия без ДЖ – у 6,8% пациенток. Число лиц женского пола репродуктивного возраста (до 45 лет) в когорте пациенток с микро- нормоцитарными

анемиями составило 4653 (58,3%). Отмечено преобладание доли женского пола с ЖДА и ЖДА+АХБ, которая составила 93,5% от выборки молодых женщин. Эту группу представляют обследованные с уровнем ФС, соответствующему группе ЖДА ( $n = 2690$ , 57,8%) и ЖДА+АХБ ( $n = 1661$ , 35,7%). ЖДА у молодых женщин встречается статистически значимо чаще по сравнению с немолодыми (коэффициент  $\chi^2 = 720,6$ ,  $p < 0,001$ ) в когорте пациентов с микро-нормоцитарными анемиями. Группа пациенток 45 лет и старше составила 3334 человека, или 41,7% от общего количества пациенток с микро-нормоцитарными анемиями. В описываемой группе ЖДА и ЖДА+АХБ выявлены у 2083 человек, что составило 62,5% этой выборки. АХБ и анемия без ДЖ у немолодых женщин встречается статистически значимо чаще, чем у молодых (коэффициент  $\chi^2 = 622,50$  и  $453,4$ ,  $p < 0,001$ , соответственно).

Число мужчин в ЖДА группе в когорте пациентов с микро-нормоцитарными анемиями составило 156 (8,7%), в ЖДА+АХБ – 403 (22,6%). Число мужчин в группе анемии без ДЖ группе (ФС  $\geq 300$  мкг/л) составило 695 (39,0%). Таким образом, большинство пациентов мужского пола с микро-нормоцитарными анемиями имеют нормальные показатели ФС. Группа с анемией без ДЖ составила 39,0%, ДЖ выявлен у 31,3%, АХБ – у 29,7% пациентов. Число молодых мужчин (до 45 лет) в когорте пациентов с микро-нормоцитарными анемиями составило 268 (15,0%). Доля мужчин с ДЖ составила 37,7% от выборки молодых мужчин. Эту группу представляют пациенты с уровнем ФС, соответствующему ЖДА ( $n = 39$ , 14,6%) и ЖДА+АХБ ( $n = 62$ , 23,1%). ЖДА у молодых мужчин встречается статистически значимо чаще по сравнению с немолодыми (коэффициент  $\chi^2 = 12,49$ ,  $p < 0,001$ ) в когорте пациентов с микро-нормоцитарными анемиями. Группа лиц мужского пола старше 45 лет составила 1516 человека, или 85,0% от общего количества с микро-нормоцитарными анемиями. В описываемой группе ДЖ выявлен у 458 человек, что составило 30,2% этой выборки. Эту группу представляют пациенты с уровнем ФС, соответствующему ЖДА ( $n = 117$ , 7,7%) и ЖДА+АХБ ( $n = 341$ , 22,5%).

Согласно критерию ( $|r_s| > 0,3$ ) определена следующая значимость показателей в прогнозировании уровня ФС в когорте пациентов с микро-нормоцитарными анемиями: возраст ( $r_s = 0,41$ ), СРБ ( $r_s = 0,47$ ), RDW-CV ( $r_s = 0,33$ ), RBC ( $r_s = 0,40$ ), MCV ( $r_s = 0,53$ ), MCH ( $r_s = 0,54$ ), MCHC ( $r_s = 0,38$ ) для лиц женского пола; СРБ ( $r_s = 0,44$ ), MCV ( $r_s = 0,34$ ), MCH ( $r_s = 0,36$ ) – для лиц мужского пола.

**Разработка и оценка диагностической точности полученных алгоритмов машинного обучения, обученных и протестированных на общей выборке пациентов.** На основе полученных данных разработан алгоритм регрессии, предназначенный для

количественного определения ФС на основе демографических и лабораторных показателей, известного уровня ФС у конкретного пациента (Таблица 1). Выявлена сильная корреляция между известным и прогнозируемым уровнем ФС ( $r = 0,86$ ,  $p < 0,05$ ).

Таблица 1 – Характеристики алгоритма регрессии в общей выборке пациентов

$R^2$	MedAE, мкг/л	MAPE, %	RMSE, мкг/л
0,71	28,6	0,72	171,1

Примечание:  $R^2$  – коэффициент детерминации; MedAE – медианная абсолютная ошибка; MAPE – средняя абсолютная ошибка в процентах; RMSE – корень из среднеквадратичной ошибки.

Коэффициент детерминации  $R^2$  достиг 0,71 (71% вариации уровня ФС успешно объясняется выбранными предикторами). MedAE, как медиана абсолютных разностей между предсказанными и фактическими значениями, составила 28,6 мкг/л на всём диапазоне определения ФС, в том числе при значительно высоких значениях ФС. Высокий показатель MAPE = 0,72 (в среднем предсказания отклоняются от реальных значений на 72% от их величины) указывает на значительные относительные ошибки, RMSE = 171 мкг/л (корень из среднеквадратичной ошибки предсказания отличается от фактического на 171 мкг/л с учетом поправок при значительных отклонениях) подтверждает значительные абсолютные ошибки, которые, вероятно, связаны с выбросами или неоднородностью набора данных. Однако, полученные результаты демонстрируют умеренную точность разработанной модели регрессии.

В Таблице 2 представлена диагностическая точность алгоритма классификации в зависимости от пола пациентов и при различных вариантах ДЖ по уровню ФС на тестовой выборке. В Таблице 3 указан фактор возраста до и после 45 лет включительно.

Таблица 2 – Характеристики алгоритма классификации с учетом пола пациентов на общей выборке пациентов

Группы по содержанию ФС (мкг/л)	АДЖ (ФС <15)		Возможный ДЖ (15 ≤ ФС <100)		ОДЖ (100 ≤ ФС <300)		ИЖ (ФС ≥300)	
	Ж	М	Ж	М	Ж	М	Ж	М
ROC AUC (OvR)*	0,88	0,94	0,75	0,74	0,78	0,65	0,85	0,76
Чувствительность	46,7	26,7	92,0	44,8	20,3	81,1	16,5	30,5
Специфичность	96,8	99,5	37,5	89,7	95,1	44,0	99,3	94,6

Примечание: AUC ROC – площадь под ROC-кривой, рассчитанная по методу One-vs-Rest (OvR); \* $p < 0,001$ .

Таким образом, прогностическая точность алгоритма классификации вариантов ДЖ по уровню ФС у лиц мужского и женского пола составила от умеренной (AUC ROC 0,65) до высокой (0,94). Наиболее приемлемо использование данного алгоритма у лиц женского пола при значениях ФС <15 и  $\geq 300$  мкг/л и у лиц мужского пола при ФС <15 мкг/л (AUC ROC >0,80).

Таблица 3 – Параметры качества алгоритма классификации в различном возрастном диапазоне на общей выборке пациентов (до и после 45 лет включительно)

Пациенты в разных группах ДЖ	АДЖ (ФС <15)		Возможный ДЖ (15 ≤ ФС <100)		ОДЖ (100 ≤ ФС <300)		ИЖ (ФС ≥300)	
	Ж	М	Ж	М	Ж	М	Ж	М
ROC AUC (OvR)* <45 лет	0,85	0,93	0,74	0,82	0,74	0,70	0,67	0,79
ROC AUC (OvR)* ≥45 лет	0,91	0,94	0,70	0,68	0,68	0,61	0,82	0,72

Примечание: AUC ROC – площадь под ROC-кривой, рассчитанная по методу One-vs-Rest (OvR); \*p <0,001.

Таким образом, прогностическая точность алгоритма классификации ДЖ по уровню ФС у лиц мужского и женского пола в двух возрастных диапазонах составила от умеренной (AUC ROC 0,61) до высокой (0,94). Наиболее точное прогнозирование вариантов ДЖ с помощью алгоритма наблюдается в группе всех пациентов с АДЖ (ФС <15 мкг/л), где значения AUC ROC достигают 0,85 – 0,94, особенно у лиц мужского пола  $\geq 45$  и <45 лет.

**Разработка и оценка диагностической точности полученных алгоритмов машинного обучения, обученных и протестированных на когорте пациентов с микро-нормоцитарными анемиями.** На основе полученных данных разработан алгоритм регрессии, предназначенный для количественного определения ФС на основе описанных выше демографических и лабораторных показателей, известного уровня ФС у конкретного пациента (Таблица 4). Выявлена сильная корреляция между известным и прогнозируемым уровнем ФС ( $r = 0,85$ ,  $p < 0,05$ ).

Таблица 4 – Характеристики алгоритма регрессии в когорте пациентов с микро-нормоцитарными анемиями

R <sup>2</sup>	MedAE, мкг/л	MAPE, %	RMSE, мкг/л
0,70	10,7	0,69	171,9



ROC AUC (OvR)* <45 лет	0,88	0,99	0,82	0,66	0,80	0,67	0,90	0,82
ROC AUC (OvR)* ≥45 лет	0,93	0,97	0,75	0,77	0,77	0,74	0,85	0,82

Таким образом, прогностическая точность алгоритма классификации ЖДС по уровню ФС у лиц мужского и женского пола в двух возрастных диапазонах составила от умеренной (AUC ROC 0,66) до высокой (0,99). Наиболее точное прогнозирование ЖДС с помощью алгоритма наблюдается в группе пациентов с ЖДА (ФС <15 мкг/л) и ИЖ (ФС ≥300 мкг/л) для лиц мужского и женского пола до и после 45 лет включительно, где значения AUC ROC достигают 0,82 – 0,99, а также для лиц женского пола с  $15 \leq \text{ФС} < 100$  и  $100 \leq \text{ФС} < 300$  (AUC ROC >0,80).

**Сопоставление полученных данных между вариантами дефицита железа в общей выборке и железодефицитными синдромами в выборке с микро- нормоцитарными анемиями.** Представленные данные свидетельствуют о том, что наличие анемии определяет большие сдвиги значений ФС как в сторону уменьшения, так и в сторону увеличения. Показатель MedAE в группах с ДЖ обладал меньшей погрешностью прогнозирования уровня ФС в когорте пациентов с анемиями по сравнению с общей выборкой пациентов, а в группах без ДЖ наоборот, то есть, использование полученного алгоритма МО в наибольшей степени рекомендуется при подозрении на ЖДА, чем при дифференциальной диагностике, предполагаемой АХБ или другой, не связанной с ДЖ анемии. Полученные результаты говорят о возможно более значимом практическом использовании данного алгоритма именно при уже выявленной анемии (в последнем случае у пациентов обоих полов показатель AUC ROC-анализа характеризовал лучшую точность, а показатели чувствительности и специфичности – лучшие диагностические характеристики, за исключением двух групп молодых мужчин (ЖДС при уровне  $15 \leq \text{ФС} < 100$  мкг/л и  $100 \leq \text{ФС} < 300$  мкг/л). Таким образом, минимальные показатели ФС по полученным данным определяют большую точность разработанных моделей, а наличие анемии и в большинстве случаев женского пола улучшает точность разработанных моделей, при этом использование классификационной модели в большинстве случаев показало лучшие результаты, чем регрессионной при всех ЖДС у лиц мужского и женского пола как моложе 45, так старше 45 лет включительно (все  $p < 0,001$ ).

**Оценка диагностической точности полученных алгоритмов машинного обучения на когорте пациентов с клиническим диагнозом железодефицитная анемия.** Результат прогнозируемого уровня ФС оказался сопоставим с измеренным в представленной выборке. Статистически значимых отличий между показателями у молодых и немолодых лиц женского

пола выявлено не было. Полученные метрики ( $R^2 = 0,61$ ,  $MedAE = 6,4$  мкг/л) свидетельствуют об умеренной, но клинически приемлемой точности алгоритма регрессии при его применении к независимой выборке с верифицированным диагнозом.  $MedAE$  в  $6,4$  мкг/л является низкой, особенно в диагностически значимом диапазоне  $15 \leq \Phi C < 100$  мкг/л, что позволяет считать модель полезным инструментом для скрининговой оценки запасов железа. Результаты показателей гемограммы соответствуют данным микро- нормоцитарной анемии и укладываются в клинический диагноз ЖДА.

Алгоритм классификации продемонстрировал стабильно умеренное качество (AUC ROC от  $0,70$  до  $0,75$ ) в клинической группе пациентов с пограничным уровнем  $\Phi C$  ( $15 \leq \Phi C < 100$  мкг/л). Небольшое снижение точности у пациентов  $\geq 45$  лет (AUC ROC  $0,70$ ) подтверждает, что на данной выборке модель работает несколько хуже. Несмотря на это, значения AUC ROC  $> 0,70$  позволяют считать модель полезным вспомогательным диагностическим инструментом и в этой группе.

Анализ случаев с наибольшими расхождениями показал, что основные причины ошибок алгоритма носят ятрогенный характер и связаны с медицинскими вмешательствами. К ним относятся: внутривенное введение препаратов железа, гемотрансфузии, назначение пероральных препаратов железа между выполнением КАК и непосредственным измерением  $\Phi C$ , причем диапазон разницы во времени в указанных случаях составил свыше двух недель. Таким образом, разработанный алгоритм демонстрирует высокую точность в стабильных клинических условиях, но его прогноз может нарушаться при динамических изменениях запасов железа вследствие терапии.

## ВЫВОДЫ

1. ДЖ ( $\Phi C < 100$  мкг/л) выявлен у подавляющего количества пациенток женского пола в общей популяции ( $75,8\%$ ) по сравнению с лицами мужского ( $31,2\%$ ),  $p < 0,001$ , причем среди женщин репродуктивного возраста его частота составила  $89,5\%$ . ИЖ ( $\Phi C > 300$  мкг/л) встречается в  $12,0\%$  от общей популяции, и практически не определяется у пациенток женского пола репродуктивного возраста ( $1,3\%$ ). Наиболее важные предикторы в прогнозе уровня  $\Phi C$  на общей выборке пациентов: возраст, СРБ, НГВ, МСV, МСН (для всех  $p < 0,001$ ).
2. ЖДА и ЖДА+АХБ ( $\Phi C < 100$  мкг/л) выявлены у подавляющего количества пациенток женского пола с микро- нормоцитарными анемиями ( $80,6\%$ ) по сравнению с лицами мужского ( $31,3\%$ ),  $p < 0,001$ , причем среди женщин репродуктивного возраста частота анемий с ДЖ составила  $93,5\%$ . Противопоказания к препаратам железа ( $\Phi C \geq 300$  мкг/л)

выявлены у 6,8% пациентов женского и у 39,0% пациентов мужского пола ( $p < 0,05$ ). Наиболее важные предикторы в прогнозе уровня ФС на выборке пациентов с микро-нормоцитарными анемиями помимо указанных в общей выборке – это RDW-CV, RBC, MCHC (для всех  $p < 0,001$ ).

3. Алгоритм регрессии, обученный и протестированный на общей выборке пациентов, получил адекватную предсказательную способность ( $R^2 = 0,71$ ). Алгоритм классификации показал высокую диагностическую точность для всех групп по содержанию ФС, максимальная значимость для группы пациентов женского пола с возможным ДЖ ( $15 \leq \text{ФС} < 100$  мкг/л) (0,75, 92,0%, 37,5%,  $p < 0,001$ , соответственно), минимальная – с ИЖ ( $\text{ФС} > 300$  мкг/л) (AUC ROC – 0,85, чувствительность – 16,5%, специфичность – 99,3%,  $p < 0,001$ ).
4. Алгоритм регрессии, обученный и протестированный на когорте пациентов с микро-нормоцитарными анемиями, получил адекватную предсказательную способность ( $R^2 = 0,70$ ). Алгоритм классификации показал высокую диагностическую точность для всех ЖДС, с такой же закономерностью, как и в общей выборке, при этом максимальная значимость для группы пациентов женского пола с ЖДА ( $\text{ФС} < 15$  мкг/л) (0,91, 82,3%, 81,7%,  $p < 0,001$ , соответственно), минимальная – с АХБ ( $100 \leq \text{ФС} < 300$  мкг/л) (AUC ROC – 0,84, чувствительность – 26,9%, специфичность – 95%,  $p < 0,001$ ).
5. Наиболее точные результаты разработанные модели показывают при наличии анемии, низких показателях ФС у пациенток женского пола. При создании регрессионной модели МО в когорте пациентов с микро-нормоцитарными анемиями результаты оказались более точными, чем в общей группе, использование классификационной модели в большинстве случаев также показало лучшие результаты при всех ЖДС у лиц мужского и женского пола как моложе 45, так старше 45 лет включительно (все  $p < 0,001$ ).
6. Алгоритм продемонстрировал удовлетворительную точность прогноза концентрации ФС при внешней валидации выборки пациентов с клиническим диагнозом ЖДА. Основными ограничениями применения полученного алгоритма является интервал (2-3 недели и более) между лабораторным исследованием КАК, СРБ и ФС и подсчетом прогностического ФС, в который проводилось то или иное лечение анемии.

### **ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ**

1. Данное исследование позволяет рекомендовать обязательное определение ФС у всех женщин репродуктивного возраста. В условиях ограниченных возможностей для проведения тотального лабораторного контроля этого показателя целесообразно или

широко внедрять полученные в работе технологии, или ориентироваться на рутинную профилактику ДЖ в указанной группе пациенток. У больных мужского пола с диагностированной микро- нормоцитарной анемией назначение препаратов железа недопустимо без предварительного обследования на содержание ФС.

2. Разработанное и апробированное ПО, предназначенное для предсказания уровня ФС, демонстрирует высокую точность и может быть рекомендовано для широкого внедрения в клиническую практику в целях проведения популяционного скрининга пациентов любого пола и возраста, что позволит минимизировать проведение дорогостоящих лабораторных анализов.
3. Полученное и протестированное в работе ПО может быть рекомендовано для дифференциальной диагностики различных вариантов ЖДС, а полученные результаты могут являться основанием для принятия решения о направлении дальнейшего обследования и назначении или отказе от лечения препаратами железа вне зависимости от пола и возраста пациента.
4. Клиническое использование созданного интеллектуального предсказательного метода вне зависимости от выбранного паттерна анализа особенно оправдано в группах пациентов женского пола с микро- нормоцитарными анемиями.
5. Алгоритм не применим, если пациент после сдачи анализов (КАК, СРБ) и до применения данного алгоритма (прогноз ФС) получил то или иное лечение по поводу анемии.

### СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

1. **Вареха Н.В.** Искусственный интеллект в лабораторной диагностике железодефицитной анемии (обзор литературы) / **Н.В. Вареха**, Н.И. Стуклов, Р.Р. Гимадиев и др. // **Клиническая лабораторная диагностика.** – 2025. – Т. 70. – № 2. – С. 102-7. – DOI: 10.51620/0869-2084-2025-70-2-102-107 [**Scopus**].
2. **Вареха Н.В.** Создание и оценка значимости прогностических моделей для определения уровня ферритина сыворотки с помощью машинного обучения в разных клинических группах / **Н.В. Вареха**, Н.И. Стуклов, Р.Р. Гимадиев и др.// **Клиническая лабораторная диагностика.** – 2025. – Т. 70. – № 3. – С. 172–81. – DOI: 10.51620/0869-2084-2025-70-3-172-181 [**Scopus**].
3. **Вареха Н.В.** Коррекция анемии препаратом железа с гепсидин-независимым механизмом всасывания у пациентов с хронической сердечной недостаточностью с сохраненной фракцией выброса левого желудочка / А.А. Гуркина, Н.И Стуклов, В.А. Кокорин и др.// **Рациональная фармакотерапия в кардиологии.** – 2025. – Т. 21. – № 1. – С. 33–38. – DOI: 10.20996/1819-6446-2025-3121 [**Scopus/Web of Science**].

4. **Вареха Н.В.** Разработка метода дифференциальной диагностики железодефицитной анемии и анемии хронических болезней на основе демографических данных и результатов рутинных лабораторных исследований с использованием технологий машинного обучения / **Н.В. Вареха, Н.И. Стуклов, К.В. Гордиенко** и др. // **Онкогематология.** – 2025. – Т. 20. – № 1. – С. 171–81. – DOI: 10.17650/1818-8346-2025-20-1-171-181 [**Scopus**].
5. **Вареха Н.В.** Программа «Ферритин 2023» по изучению популяционных показателей ферритина сыворотки у женщин в Московской агломерации (2023–2024 гг.) / **Н. И. Стуклов, М. С. Ковальчук, А. А. Гуркина** и др. // **Акушерство и гинекология: Новости. Мнения. Обучения.** - 2024. - Т. 12. - № 3 (45). - С. 94–101. – DOI: 10.33029/2303-9698-2024-12-3-94-101 [**ВАК, K1**].
6. **Вареха Н.В.** Сравнительный анализ эффективности терапии препаратом железа с гепсидин-независимым механизмом всасывания у пациентов с хронической сердечной недостаточностью / **Гуркина А.А., Стуклов Н.И., Кокорин В.А.** и др. // **Клиническая медицина.** – 2025. – Т. 103 – № 2 – С. 129-136 – DOI: 10.30629/0023-2149-2025-103-2-129-136 [**ВАК, K2**].
7. **Вареха Н.В.** Дифференциальная диагностика железодефицитной анемии и анемии хронических болезней с помощью искусственного интеллекта / **Н.В. Вареха, Н.И. Стуклов, Р.Р. Гимадиев** и др.// **Терапия.** – 2024. – Т. 10. – № 6. – С. 84. [**ВАК, K1**].

### СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

АДЖ – абсолютный дефицит железа	МАРЕ – средняя абсолютная ошибка в процентах
АХБ – анемия хронических заболеваний	МСН – среднее содержание гемоглобина в эритроците
ДЖ – дефицит железа	МСНС – средняя концентрация гемоглобина в эритроцитах
ЖДА – железодефицитная анемия	MCV – средний объем эритроцитов
ЖДС – железодефицитные синдромы	MedAE – медианная абсолютная ошибка
ИЖ – избыток железа	MON% – относительное содержание моноцитов
КАК – клинический анализ крови	NEU% – относительное содержание нейтрофилов
МО – машинное обучение	PLT – количество тромбоцитов
ОДЖ – отсутствие ДЖ	R <sup>2</sup> – коэффициент детерминации
СОЭ – скорость оседания эритроцитов	RBC – количество эритроцитов
СРБ – С-реактивный белок	RDW-CV – ширина распределения эритроцитов по объему
ФС – ферритин сыворотки	RMSE – корень из среднеквадратичной ошибки
AUC ROC – площадь под ROC кривой	r <sub>s</sub> – коэффициент корреляции Спирмена
BAS% – относительное содержание базофилов	WBC – количество лейкоцитов
EOS% – относительное содержание эозинофилов	
HCT – гематокрит	
HGB – гемоглобин	
LYM% – относительное содержание лимфоцитов	

**Вареха Николай Вячеславович (Российская Федерация)**

**Прогнозирование уровня ферритина сыворотки в общей популяции и при различных железодефицитных синдромах с использованием алгоритмов машинного обучения**

В работе представлены результаты исследования по разработке и оценке точности алгоритмов машинного обучения по прогнозированию ФС для повышения эффективности профилактики, диагностики, дифференциальной диагностики и лечения состояний, связанных с нарушением обмена железа, в различных по возрасту и полу когортах пациентов. Был проведен анализ эпидемиологии этой группы заболеваний и возможности использования полученного программного обеспечения, что позволило выделить наиболее предрасположенные к дефициту железа когорты пациентов, обосновать возможность широкого внедрения данного метода как для скрининга нарушений обмена железа, так и для дифференциальной диагностики различных вариантов железодефицитных синдромов. Алгоритм разработан на основе демографических данных (пол и возраст), данных клинического анализа крови и С-реактивного белка.

**Varekha Nikolai Vyacheslavovich (Russian Federation)**

**Prediction of serum ferritin levels in the general population and in various iron deficiency syndromes using machine learning algorithms**

This research demonstrates the results of the study on the development and accuracy assessment of machine learning algorithms used for serum ferritin prediction. It plays an important role in improving efficacy in preventing, diagnosing, differential diagnosis, and treatment of conditions associated with iron metabolism disorders in patient cohorts of varying age and gender. An analysis of the epidemiology of this group of diseases and the feasibility of using the resulting software enabled the identification of patient cohorts most susceptible to iron deficiency and the substantiation of the widespread adoption of this method for both screening for iron metabolism disorders and differential diagnosis of various iron deficiency syndromes. The algorithm was developed based on demographic data (gender and age), clinical blood count data, and C-reactive protein levels.