

ИДЕНТИФИКАЦИЯ ДИАГНОСТИЧЕСКИХ БИОМАРКЕРОВ ШИЗОФРЕНИИ НА ОСНОВЕ ИССЛЕДОВАНИЯ ПАРАМЕТРОВ ИММУНИТЕТА И СИСТЕМНОГО ВОСПАЛЕНИЯ С ПОМОЩЬЮ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ

DOI: 10.17691/stm2023.15.6.01

УДК 616.895.8–078:612.017.1

Поступила 6.10.2023 г.



И.К. Малашенкова, к.м.н., начальник лаборатории молекулярной иммунологии и вирусологии¹;
старший научный сотрудник лаборатории клинической иммунологии²;

С.А. Крынский, к.м.н., старший научный сотрудник лаборатории молекулярной иммунологии и вирусологии¹;
младший научный сотрудник отдела психических расстройств при нейродегенеративных заболеваниях
головного мозга³;

Д.П. Огурцов, к.м.н., старший научный сотрудник лаборатории молекулярной иммунологии и вирусологии¹;
научный сотрудник лаборатории клинической иммунологии²;

Н.А. Хайлов, к.м.н., старший научный сотрудник ресурсного центра молекулярной и клеточной биологии¹;

П.В. Дружинина, аспирант⁴; стажер-исследователь лаборатории биомедицински обоснованного ИИ (BIMAI-lab)⁵;

А.В. Бернштейн, д.ф.-м.н., профессор⁴;

А.В. Артемов, к.ф.-м.н., старший научный сотрудник⁴;

Г.Ш. Мамедова, младший научный сотрудник лаборатории фундаментальных методов исследования³;

Н.В. Захарова, к.м.н., руководитель лаборатории фундаментальных методов исследования³;

Г.П. Костюк, д.м.н., профессор, главный врач³;

В.Л. Ушаков, к.б.н., доцент, ведущий научный сотрудник⁶; старший научный сотрудник отдела
психических расстройств при нейродегенеративных заболеваниях головного мозга³;
старший научный сотрудник⁷;

М.Г. Шараев, старший преподаватель⁴; руководитель лаборатории биомедицински обоснованного ИИ (BIMAI-lab)⁵;
старший научный сотрудник отдела психических расстройств при нейродегенеративных заболеваниях
головного мозга³

¹Национальный исследовательский центр «Курчатовский институт», пл. Академика Курчатова, 1, Москва, 123182;

²Федеральный научно-клинический центр физико-химической медицины ФМБА России, ул. Малая Пироговская, 1а,
Москва, 119435;

³Психиатрическая клиническая больница №1 им. Н.А. Алексеева Департамента здравоохранения Москвы,
Загородное шоссе, 2, Москва, 117152;

⁴Сколковский институт науки и технологий (Сколтех), Территория Инновационного центра «Сколково»,
Большой бульвар, 30, стр. 1, Москва, 121205;

⁵The University of Sharjah, the United Arab Emirates, Sharjah, 27272;

⁶Институт перспективных исследований мозга Московского государственного университета им. М.В. Ломоносова,
Ломоносовский проспект, 27, кор. 1, Москва, 119192;

⁷Национальный исследовательский ядерный университет «МИФИ», Каширское шоссе, 31, Москва, 115409

В настоящее время показано, что существенную роль в развитии и прогрессировании шизофрении играют нарушения системного иммунитета и иммунных процессов в мозге. Тем не менее лишь единичные работы посвящены изучению возможности использования отдельных параметров иммунитета для объективизации диагноза с помощью машинного обучения. При этом к совокупности данных, полноценно отражающих системные характеристики иммунного статуса (параметры адаптивного иммунитета,

Для контактов: Малашенкова Ирина Константиновна, e-mail: malashenkova.irina@bk.ru

уровень маркеров воспаления, содержание основных цитокинов) методы машинного обучения до сих пор не применялись. С учетом комплексного характера нарушений иммунной системы при шизофрении включение широкой панели иммунологических данных в модели машинного обучения перспективно для увеличения точности классификации и выделения параметров, отражающих характерные для большинства больных иммунные расстройства.

Цель настоящей работы — исследование возможности использования иммунологических параметров для объективизации диагноза шизофрении с помощью моделей машинного обучения.

Материалы и методы. Проанализировано 17 иммунологических показателей у 63 больных шизофренией и 36 здоровых добровольцев. Методом иммуноферментного анализа определяли параметры гуморального иммунитета, системный уровень ключевых цитокинов адаптивного иммунитета, противовоспалительных и провоспалительных цитокинов, а также других маркеров воспаления. Использованные методы машинного обучения охватывали основные группы подходов обучения с учителем (supervised learning), такие как линейные модели (logistic regression), квадратичный дискриминантный анализ (quadratic discriminant analysis, QDA), методы опорных векторов (linear SVM, RBF SVM), алгоритм k-ближайших соседей (k-nearest neighbors), гауссовские процессы (gaussian process), наивный байесовский классификатор (naive bayes), деревья решений (decision tree), а также ансамблевые модели (AdaBoost, random forest, XGBoost). Для методов машинного обучения, показавших лучшее качество, проанализирована важность признаков для предсказания с лучшего фолда. Наиболее значимые признаки были отобраны по порогу квантиль 70%.

Результаты. Лучшее качество на тестовой выборке на 10 фолдах продемонстрировала ансамблевая модель AdaBoost с площадью под кривой ошибок (ROC AUC) $0,71 \pm 0,15$ и средней точностью (ACC) $0,78 \pm 0,11$. В рамках данного исследования модель AdaBoost показала хорошее качество классификации между больными шизофренией и здоровыми добровольцами (ROC AUC более 0,70) при высокой стабильности результатов (σ менее 0,2). Установлены иммунологические параметры, наиболее значимые для дифференциации между больными и добровольцами без психических заболеваний: уровень ряда маркеров системного воспаления, активации гуморального иммунитета, провоспалительных цитокинов, иммунорегуляторных цитокинов и белков, цитокинов Th1- и Th2-звена иммунитета. Впервые показана возможность различать больных шизофренией и здоровых добровольцев с точностью более 70% с помощью машинного обучения, используя только показатели иммунитета.

Результаты исследования подтверждают высокую значимость иммунной системы в патогенезе шизофрении.

Ключевые слова: шизофрения; системное воспаление; диагностические биомаркеры; иммунитет; машинное обучение; цитокины.

Как цитировать: Malashenkova I.K., Krynskiy S.A., Ogurtsov D.P., Khailov N.A., Druzhinina P.V., Bernstein A.V., Artemov A.V., Mamedova G.Sh., Zakharova N.V., Kostyuk G.P., Ushakov V.L., Sharaev M.G. Identification of diagnostic schizophrenia biomarkers based on the assessment of immune and systemic inflammation parameters using machine learning modeling. *Sovremennye tehnologii v medicine* 2023; 15(6): 5, <https://doi.org/10.17691/stm2023.15.6.01>

English

Identification of Diagnostic Schizophrenia Biomarkers Based on the Assessment of Immune and Systemic Inflammation Parameters Using Machine Learning Modeling

I.K. Malashenkova, MD, PhD, Head of the Laboratory of Molecular Immunology and Virology¹;

Senior Researcher, Laboratory of Clinical Immunology²;

S.A. Krynskiy, MD, PhD, Senior Researcher, Laboratory of Molecular Immunology and Virology¹;

Junior Researcher, Department of Mental Disorders in Neurodegenerative Brain Diseases³;

D.P. Ogurtsov, MD, PhD, Senior Researcher, Laboratory of Molecular Immunology and Virology¹;

Researcher, Laboratory of Clinical Immunology²;

N.A. Khailov, MD, PhD, Senior Researcher, Resource Center of Molecular and Cell Biology¹;

P.V. Druzhinina, PhD Student⁴; Intern Researcher, Biomedically Informed Artificial Intelligence Laboratory (BIMAI-lab)⁵;

A.V. Bernstein, DSc, Professor⁴;

A.V. Artemov, PhD, Senior Researcher⁴;

G.Sh. Mamedova, Junior Researcher, Laboratory of Fundamental Research Methods³;

N.V. Zakharova, MD, PhD, Head of the Laboratory of Fundamental Research Methods³;

G.P. Kostyuk, MD, DSc, Professor, Head Physician³;

V.L. Ushakov, PhD, Associate Professor, Leading Researcher⁶; Senior Researcher,

Department of Mental Disorders in Neurodegenerative Brain Diseases³; Senior Researcher⁷;

M.G. Sharaev, Senior Teacher⁴; Head of the Biomedically Informed Artificial Intelligence Laboratory (BIMAI-lab)⁵;

Senior Researcher, Department of Mental Disorders in Neurodegenerative Brain Diseases³

¹National Research Center "Kurchatov Institute", 1 Akademika Kurchatova Square, Moscow, 123182, Russia;

²Federal Research and Clinical Center of Physical-Chemical Medicine, Federal Medical Biological Agency of Russia,

1A Malaya Pirogovskaya St., Moscow, 119435, Russia;

³N.A. Alekseyev Psychiatric Clinical Hospital No.1, 2 Zagorodnoye Shosse, Moscow, 117152, Russia;

⁴Skolkovo Institute of Science and Technology (Skoltech), Territory of Skolkovo Innovation Center, Bldg. 1, 30 Bolshoy Boulevard, Moscow, 121205, Russia;

⁵The University of Sharjah, the United Arab Emirates, Sharjah, 27272;

⁶Institute of Perspective Brain Research, M.V. Lomonosov Moscow State University, Bldg. 1, 27 Lomonosov Avenue, Moscow, 119192, Russia;

⁷National Research Nuclear University MEPhI, 31 Kashirskoye Shosse, Moscow, 115409, Russia

Disorders of systemic immunity and immune processes in the brain have now been shown to play an essential role in the development and progression of schizophrenia. Nevertheless, only a few works were devoted to the study of some immune parameters to objectify the diagnosis by means of machine learning. At the same time, machine learning methods have not yet been applied to a set of data fully reflecting systemic characteristics of the immune status (parameters of adaptive immunity, the level of inflammatory markers, the content of major cytokines). Considering a complex nature of immune system disorders in schizophrenia, incorporation of a broad panel of immunological data into machine learning models is promising for improving classification accuracy and identifying the parameters reflecting the immune disorders typical for the majority of patients.

The aim of the study is to assess the possibility of using immunological parameters to objectify the diagnosis of schizophrenia applying machine learning models.

Materials and Methods. We have analyzed 17 immunological parameters in 63 schizophrenia patients and 36 healthy volunteers. The parameters of humoral immunity, systemic level of the key cytokines of adaptive immunity, anti-inflammatory and pro-inflammatory cytokines, and other inflammatory markers were determined by enzyme immunoassay. Applied methods of machine learning covered the main group of approaches to supervised learning such as linear models (logistic regression), quadratic discriminant analysis (QDA), support vector machine (linear SVM, RBF SVM), k-nearest neighbors algorithm, Gaussian processes, naive Bayes classifier, decision trees, and ensemble models (AdaBoost, random forest, XGBoost). The importance of features for prediction from the best fold has been analyzed for the machine learning methods, which demonstrated the best quality. The most significant features were selected using 70% quantile threshold.

Results. The AdaBoost ensemble model with ROC AUC of 0.71 ± 0.15 and average accuracy (ACC) of 0.78 ± 0.11 has demonstrated the best quality on a 10-fold cross validation test sample. Within the frameworks of the present investigation, the AdaBoost model has shown a good quality of classification between the patients with schizophrenia and healthy volunteers (ROC AUC over 0.70) at a high stability of the results (σ less than 0.2). The most important immunological parameters have been established for differentiation between the patients and healthy volunteers: the level of some systemic inflammatory markers, activation of humoral immunity, pro-inflammatory cytokines, immunoregulatory cytokines and proteins, Th1 and Th2 immunity cytokines. It was for the first time that the possibility of differentiating schizophrenia patients from healthy volunteers was shown with the accuracy of more than 70% with the help of machine learning using only immune parameters.

The results of this investigation confirm a high importance of the immune system in the pathogenesis of schizophrenia.

Key words: schizophrenia; systemic inflammation; diagnostic biomarkers; immunity; machine learning; cytokines.

Введение

Шизофрения — тяжелое хроническое психическое заболевание с расстройствами когнитивных функций, эмоциональными и психомоторными нарушениями [1]. При шизофрении наблюдается высокая частота инвалидизации и сокращается продолжительность жизни [2]. Для уменьшения риска неблагоприятного течения заболевания необходимы ранняя диагностика и своевременное назначение терапии. Важное направление исследований для совершенствования диагностики шизофрении — разработка подходов к объективизации диагноза с помощью моделей машинного обучения (machine learning, ML) на основе клинически значимых биомаркеров [3, 4]. Биомаркеры в клинической практике важны для определения нескольких факторов: уточнения диагноза, верификации стадии заболевания, выбора оптимального плана терапии, долгосрочного прогноза заболевания. Однако для большинства найденных к настоящему времени биомаркеров шизофрении их клиническая ценность (определяемая

показателями чувствительности, специфичности, прогностической значимости и т.п.) не установлена на требуемом доказательном уровне, в силу чего использование биомаркеров не предусмотрено в существующих клинических рекомендациях [5].

Большой объем, высокая размерность и гетерогенность мультимодальных лабораторных данных затрудняют интеграцию всех доступных модальностей в рамках одного исследования, поэтому в сообществе клиницистов растет интерес к современным подходам объединения гетерогенных данных, к методам ML и глубокого обучения (deep learning, DL) [6].

В настоящее время доказано, что существенную роль в развитии и прогрессировании шизофрении играют нарушения системного иммунитета, а также иммунные процессы в центральной нервной системе (ЦНС) [7–9]. Тем не менее лишь единичные работы посвящены изучению возможности использования отдельных параметров иммунитета для объективизации диагноза шизофрении с помощью методов ML [10]. При этом к совокупности данных, полноценно

отражающих системные характеристики иммунного статуса (параметры адаптивного иммунитета, уровень маркеров воспаления, содержание основных цитокинов), методы ML до сих пор не применялись [11–13]. С учетом комплексного характера нарушений иммунной системы при шизофрении включение в модели ML широкой панели иммунологических данных перспективно для увеличения точности классификации заболевания и выделения параметров, изменение которых наиболее характерно для большинства больных. С практической точки зрения результаты анализа данных методом ML могут позволить разработать новые панели маркеров, имеющих диагностическое/прогностическое значение, а также стать основой для создания прототипа системы поддержки принятия врачебных решений на базе отобранных признаков.

Цель настоящей работы — исследование возможности использования иммунологических параметров для объективизации диагноза шизофрении с помощью моделей машинного обучения.

Материалы и методы

Клинические данные, представленные в настоящей работе, являются частью исследовательской программы «Молекулярные и нейрофизиологиче-

ские маркеры эндогенных заболеваний», проводимой в Психиатрической клинической больнице №1 им. Н.А. Алексеева Департамента здравоохранения Москвы и одобренной Независимым междисциплинарным комитетом по этической экспертизе клинических исследований (протокол №12 от 14.07.2017 г.). Некоторые результаты этого проекта опубликованы ранее [14–16].

Пациенты. Обследовано 63 пациента (36 мужчин и 27 женщин, средний возраст — 29 ± 3 года) из числа госпитализированных в Психиатрическую клиническую больницу №1 им. Н.А. Алексеева с диагнозами расстройств шизофренического спектра (F20 и F25 по МКБ-10). Все испытуемые дали письменное информированное согласие после полного описания процедур исследования согласно Хельсинкской декларации.

Критерии включения: соответствие состояния пациента критериям шизофрении по МКБ-10 и DSM-5 (Диагностическое и статистическое руководство по психическим расстройствам, 5-е издание); критика к болезненному состоянию; сохранение памяти о психотических симптомах; информированное согласие на участие в исследовании.

Критерии исключения: шизоаффективные и аффективные расстройства; органические заболевания головного мозга; тяжелые соматические и/или неврологические состояния, потенциально влияющие на физиологию или структуру мозга; признаки злоупотребления психоактивными веществами; острые хронические соматические и инфекционно-воспалительные заболевания или их обострение.

Группу контроля составили 36 добровольцев, сопоставимых с пациентами по возрасту и полу (21 мужчина и 15 женщин, средний возраст — 30 ± 2 года) и не состоящих с пациентами в родстве. Добровольцы и больные шизофренией были обследованы по единому протоколу.

Дизайн исследования: кросс-секционное, наблюдательное, случай-контроль. Диагностику состояния больных проводили в течение двух дней. Перед взятием биоматериала определяли клиническую картину психоза и уточняли объективные сведения (от родственников и с помощью медицинских карт). За сутки до взятия биоматериала выполняли психометрическую оценку состояния с применением шкалы позитивных и негативных симптомов шизофрении (PANSS) [17] (табл. 1). Клиническое обследование проводили два опытных психиатра с привлечением всех необходимых данных (опрос родственников, анализ медицинских карт, результаты физикальных и лабораторных тестов и т.д.). Психическое состояние на момент сканирования характеризовалось критичностью к перенесенному психозу при сохранении памяти с возможностью изложить субъективный анамнез развития заболевания (с детализацией фабулы бреда и оценкой аффективного состояния в тот период).

Иммунологические исследования. Проводили анализ 17 иммунологических показателей. Они вклю-

Т а б л и ц а 1

Характеристики исследуемых больных шизофренией (n=63)

Показатели	Пациенты (M±σ)
Средний возраст, лет	28,7±7,7
Средний возраст начала болезни, лет	18,0±4,2
Средний возраст манифестации заболевания, лет	21,3±4,0
Длительность болезни от продрома, лет	9,6±7,8
Длительность болезни от манифестации, лет	6,5±6,0
PANSS общее, баллы	94,1±17,1
PANSS P, баллы:	17,3±6,4
P1 (бред)	2,9±1,2
P2 (расстройства мышления)	3,1±1,3
P3 (галлюцинации)	2,5±1,8
P4 (возбуждение)	1,7±1,1
P5 (идеи величия)	1,5±0,8
P6 (подозрительность)	3,5±1,4
P7 (враждебность)	2,2±1,1
PANSS N, баллы	27,5±7,4
PANSS G, баллы	49,3±8,7
NSA-4, баллы	20,4±4,6

Примечания: NSA-4 — шкала оценки негативных симптомов; PANSS — шкала позитивных и негативных симптомов шизофрении; PANSS P — тяжесть продуктивной симптоматики; PANSS N — тяжесть негативной симптоматики; PANSS G — выраженность других психических нарушений по общей психопатологической шкале.

чали параметры гуморального иммунитета (IgA, IgM и IgG); системный уровень ключевых провоспалительных и противовоспалительных цитокинов, а также других маркеров воспаления (С-реактивный белок (CRP), кортизол, циркулирующие иммунные комплексы (ЦИК)); уровень нейротрофического фактора мозгового происхождения (BDNF), определявшийся методом иммуноферментного анализа с использованием наборов реактивов производства ООО «ХЕМА» (Россия), ООО «Цитокин» (Россия), R&D Systems (США).

Статистическая обработка данных. Для статистической обработки результатов применяли программное обеспечение Python 3.9.0 (Python Software Foundation) в виде библиотек NumPy, Pandas, SkLearn, где также были взяты методы ML.

В данном исследовании анализировали количественные непрерывные переменные. Анализ переменных в динамике не проводили, так как исследование носило кросс-секционный характер. Изучали две независимые выборки (больные шизофренией и здоровые добровольцы).

В рамках первичного анализа данных нормальность распределения оценивали по критерию Шапиро–Уилка. Клинические данные представляли в виде средних со стандартным отклонением ($M \pm \sigma$), иммунологические параметры — в виде медиан и 25-го и 75-го перцентилей ($Me [25; 75]$).

Для оценки качества классификации использовали метрики «точность» (accuracy, ACC) и «площадь под кривой ошибок» (area under the ROC curve, ROC AUC). Точность измеряет долю верно предсказанных значений от общего числа ответов, и она выше для моделей, дающих более правильные прогнозы. Однако для наборов данных, где классы представлены неравномерно, точность может быть неоптимальной метрикой. Метрика ROC AUC измеряет площадь под кривой ROC, отображая способность модели различать объекты, принадлежащие к двум классам, при различных порогах классификации. ROC AUC может принимать значения в диапазоне от 0 до 1, причем более высокое значение указывает на лучшую производительность модели. В данном исследовании, поскольку число больных шизофренией и здоровых добровольцев было различным, в качестве основной метрики качества моделей использовали ROC AUC. Пороговым уровнем ROC AUC было значение 0,70, что соответствует хорошему качеству классификации.

Для оценки стабильности полученных моделей по каждой из указанных метрик (ACC и ROC AUC) по всем фолдам рассчитывали стандартное отклонение (σ). Более низкое σ указывает на более стабильную работу модели, при этом малым считается σ менее 0,2, умеренным — от 0,2 до 0,5. В рамках данного исследования стабильными считали модели, имевшие σ ACC и ROC AUC менее 0,2.

Методы машинного обучения. Используемые методы ML охватывали основные группы подходов обучения с учителем (supervised learning), такие как

линейные модели (logistic regression), квадратичный дискриминантный анализ (quadratic discriminant analysis, QDA), методы опорных векторов (linear SVM, RBF SVM), алгоритм k-ближайших соседей (k-nearest neighbors), гауссовские процессы (gaussian process), наивный байесовский классификатор (naive bayes), деревья решений (decision tree), а также ансамблевые модели (AdaBoost, random forest, XGBoost). Эти методы охватывают широкий спектр подходов в ML, каждый из которых имеет свои сильные стороны. Рассмотрение различных групп подходов позволило установить, какие из них лучше всего справляются с данной задачей. Говоря более подробно о семействе ансамблевых подходов, следует отметить, что в их основе лежит идея комбинации нескольких классификаторов в один сильный. Например, каждый слабый классификатор в AdaBoost (adaptive boosting) может быть любым алгоритмом обучения (принимающим значения отрицательных и положительных весов). На каждой итерации обучения AdaBoost увеличивает веса неправильно классифицированных примеров, чтобы следующий классификатор был более сфокусирован на тех примерах, на которых предыдущие классификаторы сделали ошибки. После обучения все слабые классификаторы объединяются с использованием взвешивания голосов, где каждому классификатору присваивается вес, основанный на его точности при классификации, что позволяет получить сильный классификатор, представляющий собой ансамбль слабых классификаторов. Когда итоговый классификатор используется для категоризации новых примеров, он учитывает голос всех слабых классификаторов в соответствии с их весами, что позволяет достичь более точной и устойчивой классификации.

Для каждой модели ML можно оценить релевантность или вклад различных признаков относительно предсказания, т.е. количественно оценить важность каждого признака (feature importance) в процессе принятия решения моделью. Важность признака можно рассчитать с использованием различных методов, таких как важность перестановки, важность Джини (Gini impurity) или величина коэффициента важности, который позволяет определить воздействие каждой отдельной переменной на предсказываемую переменную и оценить их роль в моделировании. Результат расчета коэффициентов важности может быть представлен в виде таблицы, позволяющей проанализировать значимость каждой переменной для моделирования. Более высокие значения указывают на большую важность. Хотя важность признаков может дать представление об относительном вкладе каждого из них, данный параметр следует интерпретировать с осторожностью. Это не обязательно подразумевает причинно-следственную связь или прямую связь между признаками. Более того, интерпретация важности функции зависит от конкретной модели и контекста, и ее следует рассматривать наряду с другими мерами оценки и знаниями предметной области для

формирования всестороннего понимания. Таким образом, при использовании параметра важности функций можно оценить релевантность функций в модели ML. Данный параметр позволяет обеспечить интерпретируемость, однако его следует использовать как часть более широкого анализа, а не как единственную основу для вынесения выводов.

Процесс обучения и оценки качества моделей проводили при помощи 10-кратной стратифицированной перекрестной проверки (StratifiedKFold). В итоге было создано 10 моделей для каждого метода ML, качество каждой модели определяли на тестовой части (20% данных). Таким образом, мы получили выборку из десяти значений для каждой метрики качества для выбранного набора гиперпараметров. Затем для каждой метрики качества находили среднее и стандартное отклонение. Для методов ML, показавших лучшее качество, была проанализирована важность признаков для предсказания с лучшего фолда. Наиболее значимые признаки были отобраны по порогу квантиль 70%.

Результаты

В ходе экспериментов обнаружено, что лучшее качество на тестовой выборке на 10 фолдах продемонстрировала ансамблевая модель AdaBoost с площадью под кривой ошибок (ROC AUC) $0,71 \pm 0,15$ и

Таблица 2

Результаты объективизации диагноза шизофрении с помощью моделей машинного обучения

№ эксперимента	Методы	Бинарная классификация	
		Средняя ROC AUC (σ)	Средняя ACC (σ)
1	AdaBoost	0,71 (0,15)	0,78 (0,11)
2	Decision tree	0,56 (0,25)	0,60 (0,21)
3	Gaussian process	0,52 (0,12)	0,66 (0,08)
4	Linear SVM	0,49 (0,03)	0,67 (0,05)
5	Logistic regression	0,62 (0,21)	0,70 (0,16)
6	Naive bayes	0,58 (0,20)	0,54 (0,19)
7	k-nearest neighbors	0,63 (0,22)	0,67 (0,18)
8	Neural networks	0,58 (0,14)	0,67 (0,09)
9	Quadratic discriminant analysis	0,62 (0,15)	0,70 (0,13)
10	RBF SVM	0,50 (0,0)	0,68 (0,03)
11	Random forest	0,54 (0,17)	0,67 (0,12)
12	XGBoost classifier	0,68 (0,21)	0,75 (0,19)

Примечания: метрики ROC AUC, ACC подсчитаны как среднее по тестовому набору данных на кросс-валидации с 10 фолдами; наиболее точные результаты классификации отмечены жирным шрифтом.

средней точностью (ACC) $0,78 \pm 0,11$. В рамках данного исследования модель AdaBoost показала хорошее качество классификации между больными шизофренией и здоровыми добровольцами (ROC AUC более 0,70) при высокой стабильности результатов (σ менее 0,2) (табл. 2).

В табл. 3 представлены результаты анализа признаков по степени их важности для точности классификации с помощью модели AdaBoost.

В табл. 4 показано ранжирование признаков по их важности для классификации с помощью модели XGBoost classifier, имевшей вторые по величине ROC AUC и ACC, высокую или умеренную стабильность результатов классификации (ROC AUC — $0,68 \pm 0,21$; ACC — $0,75 \pm 0,19$).

Для повышения наглядности полученных данных также представлено ранжирование признаков по их важности для классификации с помощью модели логистической регрессии, которая отличается большей простотой и наглядностью по сравнению с AdaBoost и XGBoost classifier и имеет третье по величине значение ROC AUC с высокой или умеренной стабиль-

Таблица 3

Ранжирование важности признаков модели AdaBoost с порогом по квантилю 70% — 0,084

№	Признаки	Коэффициент важности
1	IL-8	0,12
2	ЦИК	0,1
3	IL-4	0,1
4	IFN- γ	0,1
5	TNF- α	0,1
6	Кортизол	0,08

Примечания: IFN- γ — интерферон- γ , IL — интерлейкин, TNF- α — tumor necrosis factor α (фактор некроза опухоли α), ЦИК — циркулирующие иммунные комплексы.

Таблица 4

Ранжирование важности признаков модели XGBoost classifier с порогом по квантилю 70% — 0,081

№	Признаки	Коэффициент важности
1	IL-1 β	0,128
2	IL-10	0,1
3	Кортизол	0,089
4	ЦИК	0,088
5	TNF- α	0,088

Примечания: IL — интерлейкин, TNF- α — tumor necrosis factor α (фактор некроза опухоли α), ЦИК — циркулирующие иммунные комплексы.

Таблица 5
Ранжирование важности признаков модели logistic regression с порогом по квантилю 70% — 0,303

№	Признаки	Коэффициент важности
1	IL-1RA	1,208
2	IFN- γ	0,889
3	ЦИК	0,678
4	CRP	0,574
5	IgM	0,336

Примечания: CRP — С-реактивный белок, Ig — иммуноглобулин, IFN- γ — интерферон- γ , IL-1RA — рецепторный антагонист интерлейкина 1 β , ЦИК — циркулирующие иммунные комплексы.

ностью результатов (ROC AUC — 0,62 \pm 0,21, ACC — 0,70 \pm 0,16) (см. табл. 2; табл. 5).

Таким образом, ранжирование по коэффициенту важности признаков, релевантных для трех моделей, имевших наибольшую предсказательную силу, позволило установить параметры, наиболее значимые для различения больных шизофренией и здоровых добровольцев.

Обсуждение

В результате данной работы впервые показана возможность диагностировать шизофрению с точностью более 70%, используя модели машинного обучения, построенные на основе совокупности только показателей состояния иммунной системы. Указанный уровень точности подтверждения диагноза был получен, несмотря на то, что работа на данном этапе не предусматривала включения в модели машинного обучения результатов клинического обследования, в отличие, например, от работы [10], где использовались как отдельные иммунологические данные, так и результаты нейropsychологического тестирования.

В результате работы впервые выявлен комплекс параметров естественного и адаптивного иммунитета, наиболее важных для классификации больных и лиц без психических заболеваний: уровень ряда маркеров системного воспаления и активации гуморального иммунитета (CRP, кортизол, ЦИК), провоспалительных цитокинов (IL-1 β , IL-8, TNF- α), иммунорегуляторных белков (IL-1RA), цитокинов Th1-звена иммунитета (IFN- γ), цитокинов Th2-звена иммунитета (IL-4) и иммунорегуляторных цитокинов (IL-10).

Обращает на себя внимание важная роль в построенных моделях машинного обучения показателей системного воспаления и уровня провоспалительных цитокинов. Данные литературы и наших исследований подтверждают, что у пациентов с шизофренией выявляются признаки системного воспаления, которые наиболее выражены при первом эпизоде заболевания и

при его обострениях. Так, при обострении болезни в крови больных повышаются уровни провоспалительных цитокинов IL-8 и TNF- α , а также белка острой фазы CRP [18–20]. Показано также, что высокий уровень CRP ассоциирован с более тяжелым течением психоза при шизофрении и последующим снижением когнитивных функций [18, 21]. Определенный уровень иммунновоспалительной активации, по нашим данным, сохраняется и в период медикаментозной ремиссии. Однако до сих пор не изучалось применение методов машинного обучения для подтверждения диагноза шизофрении с помощью оценки уровня маркеров системного воспаления [11, 12]. В то же время можно выделить ряд работ, в которых установлены взаимосвязи между уровнем отдельных маркеров системного воспаления и клиническими характеристиками шизофрении, включая выраженность нарушений когнитивных функций (TNF- α , IL-2, IL-6, IL-8, кортизол), нейрокогнитивного дефекта (IL-1 β , sIL-1RA, TNF- α), острый или хронический характер течения заболевания (TNF- α , IL-2, IL-6, IL-8, кортизол) [22, 23].

В запуске и поддержании системного воспалительного ответа, пирогенных реакций, а также функций адаптивного иммунитета важную роль играют IL-1 β , IL-8, TNF- α — цитокины врожденного иммунитета [24]. Повышение их уровня в ЦНС, которое может наблюдаться, в частности, при нейроинфекциях, аутоиммунных заболеваниях, психических заболеваниях, оказывает повреждающее действие на нейроны и клетки глии, способствует нейродегенерации (см., например, обзор [25]). Согласно данным метаанализа [18], уровень цитокинов IL-6, TNF- α и рецепторного антагониста IL-1RA при острой шизофрении повышен. При хронической шизофрении в этом же метаанализе показано повышение уровня IL-1 β и IL-6. Согласно нашим данным, повышенный уровень IL-8 отмечается у пациентов с шизофренией независимо от выраженности клинической симптоматики [26].

Важным цитокином, активирующим функции CD3+CD4+ Т-хелперов 1-го типа, CD3+CD8+ цитотоксических Т-клеток и CD3–CD16+CD56+ NK-клеток, которые обеспечивают противовирусную защиту, а также стимулирующим презентацию антигенов лимфоцитам клетками врожденного иммунитета, является IFN- γ . В условиях воспаления IFN- γ активно продуцируется Т-лимфоцитами и NK-клетками. По нашим результатам и по данным литературы, его содержание при шизофрении повышается, отражая комплексную активацию механизмов адаптивного и врожденного иммунитета у больных [26, 27]. В рамках данной работы показано, что IFN- γ является одним из наиболее значимых показателей для подтверждения диагноза шизофрении с помощью моделей машинного обучения.

К числу диагностически значимых маркеров, выявленных в результате данной работы, относятся также иммунологические параметры, отражающие активацию Th2-звена иммунитета и иммунорегуляторных

механизмов: уровень цитокина IL-4, иммунорегуляторного белка IL-1RA, а также ЦИК. Согласно литературным данным, активация врожденного иммунного ответа у больных шизофренией сопровождается активацией Th2-звена адаптивного иммунитета. Признаки активации Th2-иммунного ответа при шизофрении включают увеличение уровня цитокинов IL-4 и IL-10 в сыворотке крови, снижение отношения уровней Th1-цитокинов к уровням Th2-цитокинов (IFN- γ /IL-4, IFN- γ /IL-10, IL-2/IL-4, TNF- α /IL-4) [28]. S.M. de Campos-Carli с соавт. [29] обнаружили у больных связь между выраженностью когнитивных нарушений и повышением уровня в крови Th2-цитокина IL-33 [29]. Активация Th2-ответа у пациентов с шизофренией может участвовать в нарушениях обмена нейротрансмиттеров в ЦНС, вовлеченных в патогенез негативной симптоматики при шизофрении [30]. Ранее нами было впервые установлено, что иммунологический профиль, характеризующийся повышением содержания IL-10 при умеренных признаках системного воспаления, ассоциирован у больных шизофренией с наличием выраженных негативных симптомов [26].

Таким образом, выявленный комплекс иммунологических показателей (IL-8, ЦИК, IL-4, IFN- γ , TNF- α , кортизол) является значимым для диагностики шизофрении. Следует отметить, что данные параметры ассоциированы с важными клиническими характеристиками заболевания. Литературные данные указывают на их участие в патогенезе шизофрении.

Определение этих иммуновоспалительных биомаркеров предполагает меньшие временные, экономические и организационные затраты, чем нейровизуализация и исследование спинномозговой жидкости. Наша работа показала перспективность использования иммунологических параметров в клинической практике для объективизации диагноза шизофрении.

Заключение

В результате данной работы впервые с помощью методов машинного обучения показана возможность подтверждения диагноза шизофрении с точностью более 70% на основании оценки только параметров системного воспаления, врожденного и адаптивного иммунитета. Результаты являются подтверждением патогенетической значимости состояния иммунной системы и иммуновоспалительного синдрома при шизофрении.

В рамках следующего этапа исследования планируется анализ широкого спектра иммунологических параметров в составе мультимодальных данных с использованием методов машинного обучения для создания интерпретируемых моделей, которые могут быть использованы для выявления клинико-иммунологических вариантов психозов.

Источники финансирования. Работа проведена в рамках выполнения государственного задания НИЦ

«Курчатовский Институт» и частично — за счет гранта Российского научного фонда №21-71-10136, <https://rscf.ru/project/21-71-10136/>.

Конфликт интересов. Конфликт интересов, связанный с данным исследованием, отсутствует.

Литература/References

1. Jääskeläinen E., Juola P., Hirvonen N., McGrath J.J., Saha S., Isohanni M., Veijola J., Miettunen J. A systematic review and meta-analysis of recovery in schizophrenia. *Schizophr Bull* 2013; 39(6): 1296–1306, <https://doi.org/10.1093/schbul/sbs130>.
2. Nordentoft M., Wahlbeck K., Hällgren J., Westman J., Osby U., Alinaghizadeh H., Gissler M., Laursen T.M. Excess mortality, causes of death and life expectancy in 270,770 patients with recent onset of mental disorders in Denmark, Finland and Sweden. *PLoS One* 2013; 8(1): e55176, <https://doi.org/10.1371/journal.pone.0055176>.
3. Correll C.U., Galling B., Pawar A., Krivko A., Bonetto C., Ruggeri M., Craig T.J., Nordentoft M., Srihari V.H., Guloksuz S., Hui C.L.M., Chen E.Y.H., Valencia M., Juarez F., Robinson D.G., Schooler N.R., Brunette M.F., Mueser K.T., Rosenheck R.A., Marcy P., Addington J., Estroff S.E., Robinson J., Penn D., Severe J.B., Kane J.M. Comparison of early intervention services vs treatment as usual for early-phase psychosis: a systematic review, meta-analysis, and meta-regression. *JAMA Psychiatry* 2018; 75(6): 555–565, <https://doi.org/10.1001/jamapsychiatry.2018.0623>.
4. Ferrara M., Guloksuz S., Mathis W.S., Li F., Lin I.H., Syed S., Gallagher K., Shah J., Kline E., Tek C., Keshavan M., Srihari V.H. First help-seeking attempt before and after psychosis onset: measures of delay and aversive pathways to care. *Soc Psychiatry Psychiatr Epidemiol* 2021; 56(8): 1359–1369, <https://doi.org/10.1007/s00127-021-02090-0>.
5. Mirzakhani H., Singh F., Cadenhead K.S. Biomarkers in psychosis: an approach to early identification and individualized treatment. *Biomark Med* 2014; 8(1): 51–57, <https://doi.org/10.2217/bmm.13.134>.
6. Ellis J.K., Walker E.F., Goldsmith D.R. Selective review of neuroimaging findings in youth at clinical high risk for psychosis: on the path to biomarkers for conversion. *Front Psychiatry* 2020; 11: 567534, <https://doi.org/10.3389/fpsy.2020.567534>.
7. van Mierlo H.C., Schot A., Boks M.P.M., de Witte L.D. The association between schizophrenia and the immune system: review of the evidence from unbiased 'omic-studies'. *Schizophr Res* 2020; 217: 114–123, <https://doi.org/10.1016/j.schres.2019.05.028>.
8. Birnbaum R., Weinberger D.R. A genetics perspective on the role of the (neuro)immune system in schizophrenia. *Schizophr Res* 2020; 217: 105–113, <https://doi.org/10.1016/j.schres.2019.02.005>.
9. Sun H.L., Bai W., Li X.H., Huang H., Cui X.L., Cheung T., Su Z.H., Yuan Z., Ng C.H., Xiang Y.T. Schizophrenia and inflammation research: a bibliometric analysis. *Front Immunol* 2022; 13: 907851, <https://doi.org/10.3389/fimmu.2022.907851>.
10. Fernandes B.S., Karmakar C., Tamouza R., Tran T., Yearwood J., Hamdani N., Laouamri H., Richard J.R., Yolken R., Berk M., Venkatesh S., Leboyer M. Precision psychiatry with immunological and cognitive biomarkers: a

multi-domain prediction for the diagnosis of bipolar disorder or schizophrenia using machine learning. *Transl Psychiatry* 2020; 10(1): 162, <https://doi.org/10.1038/s41398-020-0836-4>.

11. Ferrara M., Franchini G., Funaro M., Cutroni M., Valier B., Toffanin T., Palagini L., Zerbinati L., Folesani F., Murri M.B., Caruso R., Grassi L. Machine learning and non-affective psychosis: identification, differential diagnosis, and treatment. *Curr Psychiatry Rep* 2022; 24(12): 925–936, <https://doi.org/10.1007/s11920-022-01399-0>.

12. Sharaev M.G., Malashenkova I.K., Maslennikova A.V., Zakharova N.V., Bemstein A.V., Burnaev E.V., Mamedova G.S., Krynskiy S.A., Ogurtsov D.P., Kondrateva E.A., Druzhinina P.V., Zubrikhina M.O., Arkhipov A.Yu., Strelets V.B., Ushakov V.L. Diagnosis of schizophrenia based on the data of various modalities: biomarkers and machine learning techniques (review). *Sovremennye tehnologii v medicine* 2022; 14(5): 53, <https://doi.org/10.17691/stm2022.14.5.06>.

13. Малашенкова И.К., Крынский С.А., Огурцов Д.П., Мамошина М.В., Захарова Н.В., Ушаков В.Л., Величковский Б.М., Дидковский Н.А. Роль иммунной системы в патогенезе шизофрении. *Журнал неврологии и психиатрии им. С.С. Корсакова* 2018; 118(12): 72–80, <https://doi.org/10.17116/jnevro201811812172>.

Malashenkova I.K., Krynskiy S.A., Ogurtsov D.P., Mamoshina M.V., Zakharova N.V., Ushakov V.L., Velichkovsky B.M., Didkovsky N.A. A role of the immune system in the pathogenesis of schizophrenia. *Zhurnal neurologii i psikiatrii im. S.S. Korsakova* 2018; 118(12): 72–80, <https://doi.org/10.17116/jnevro201811812172>.

14. Corponi F., Zorkina Y., Stahl D., Murru A., Vieta E., Serretti A., Morozova A., Reznik A., Kostyuk G., Chekhonin V.P. Frontal lobes dysfunction across clinical clusters of acute schizophrenia. *Rev Psiquiatr Salud Ment* 2021, <https://doi.org/10.1016/j.rpsm.2021.12.002>.

15. Morozova A., Zorkina Y., Abramova O., Pavlova O., Pavlov K., Soloveva K., Volkova M., Alekseeva P., Andryshchenko A., Kostyuk G., Gurina O., Chekhonin V. Neurobiological highlights of cognitive impairment in psychiatric disorders. *Int J Mol Sci* 2022; 23(3): 1217, <https://doi.org/10.3390/ijms23031217>.

16. Kartashov S.I., Vartanov A.V., Zakharova N.V., Arkhipov A.Yu., Strelets V.B., Maslennikova A.V., Ushakov V.L. Tractographic studies in the human brain at the norm and the paranoid schizophrenia. *Procedia Comput Sci* 2021; 190: 434–438, <https://doi.org/10.1016/j.procs.2021.06.051>.

17. Kay S.R., Fiszbein A., Opler L.A. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* 1987; 13(2): 261–276, <https://doi.org/10.1093/schbul/13.2.261>.

18. Goldsmith D.R., Rapaport M.H., Miller B.J. A meta-analysis of blood cytokine network alterations in psychiatric patients: comparisons between schizophrenia, bipolar disorder and depression. *Mol Psychiatry*. 2016; 21(12): 1696–1709, <https://doi.org/10.1038/mp.2016.3>.

19. Malashenkova I.K., Krynskiy S.A., Ogurtsov D.P., Hailov N.A., Zakharova N.V., Bravve L.V., Kaydan M.A., Chekulaeva E.I., Andreyuk D.S., Ushakov V.L., Didkovsky N.A., Kostyuk G.P. Immunoinflammatory profile in patients with episodic and continuous paranoid schizophrenia. *Consortium Psychiatricum* 2021; 2(1): 19–31, <https://doi.org/10.17816/cp66>.

20. Ushakov V.L., Malashenkova I.K., Kostyuk G.P., Zakharova N.V., Krynskiy S.A., Kartashov S.I., Ogurtsov D.P., Bravve L.V., Kaydan M.A., Hailov N.A., Chekulaeva E.I., Didkovsky N.A. The relationship between inflammation, cognitive impairments, and neuroimaging data in schizophrenia. *Neurosci Behav Physiol* 2021; 51(7): 873–881.

21. Johnsen E., Fathian F., Kroken R.A., Steen V.M., Jørgensen H.A., Gjestad R., Løberg E.M. The serum level of C-reactive protein (CRP) is associated with cognitive performance in acute phase psychosis. *BMC Psychiatry*. 2016; 16: 60, <https://doi.org/10.1186/s12888-016-0769-x>.

22. Zhang Q., He H., Cao B., Gao R., Jiang L., Zhang X., Dai J. Analysis of cognitive impairment in schizophrenia based on machine learning: Interaction between psychological stress and immune system. *Neurosci Lett* 2021; 760: 136084, <https://doi.org/10.1016/j.neulet.2021.136084>.

23. Al-Hakeim H.K., Almulla A.F., Al-Dujaili A.H., Maes M. Construction of a neuro-immune-cognitive pathway-phenotype underpinning the phenome of deficit schizophrenia. *Curr Top Med Chem* 2020; 20(9): 747–758, <https://doi.org/10.2174/1568026620666200128143948>.

24. Tanaka T., Narazaki M., Kishimoto T. IL-6 in inflammation, immunity, and disease. *Cold Spring Harb Perspect Biol* 2014; 6(10): a016295, <https://doi.org/10.1101/cshperspect.a016295>.

25. Малашенкова И.К., Крынский С.А., Хайлов Н.А., Казанова Г.В., Величковский Б.Б., Дидковский Н.А. Роль цитокинов в консолидации памяти. *Успехи современной биологии* 2015; 135(5): 419–436.

Malashenkova I.K., Krynskiy S.A., Khailov N.A., Kazanova G.V., Velichkovsky B.B., Didkovsky N.A. The role of cytokines in memory consolidation. *Uspekhi sovremennoy biologii* 2015; 135(5): 419–436.

26. Malashenkova I.K., Ushakov V.L., Zakharova N.V., Krynskiy S.A., Ogurtsov D.P., Hailov N.A., Chekulaeva E.I., Ratushnyy A.Y., Kartashov S.I., Kostyuk G.P., Didkovsky N.A. Neuro-immune aspects of schizophrenia with severe negative symptoms: new diagnostic markers of disease phenotype. *Sovremennye tehnologii v medicine* 2021; 13(6): 24, <https://doi.org/10.17691/stm2021.13.6.03>.

27. Momtazmanesh S., Zare-Shahabadi A., Rezaei N. Cytokine alterations in schizophrenia: an updated review. *Front Psychiatry* 2019; 10: 892, <https://doi.org/10.3389/fpsy.2019.00892>.

28. Chiang S.S.W., Riedel M., Schwarz M., Mueller N. Is T-helper type 2 shift schizophrenia-specific? Primary results from a comparison of related psychiatric disorders and healthy controls. *Psychiatry Clin Neurosci* 2013; 67(4): 228–236, <https://doi.org/10.1111/pcn.12040>.

29. de Campos-Carli S.M., Miranda A.S., Dias I.C., de Oliveira A., Cruz B.F., Vieira É.L., Rocha N.P., Barbosa I.G., Salgado J.V., Teixeira A.L. Serum levels of interleukin-33 and its soluble form receptor (sST2) are associated with cognitive performance in patients with schizophrenia. *Compr Psychiatry* 2017; 74: 96–101, <https://doi.org/10.1016/j.comppsy.2017.01.008>.

30. Musso T., Gusella G.L., Brooks A., Longo D.L., Varesio L. Interleukin-4 inhibits indoleamine 2,3-dioxygenase expression in human monocytes. *Blood* 1994; 83(5): 1408–1411, <https://doi.org/10.1182/blood.V83.5.1408.1408>.