

©Андросова Л.В. и др., 2024

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

УДК 616.89; 571.27

<https://doi.org/10.30629/2618-6667-2024-22-1-6-14>

## Диагностика и оценка тяжести болезни Альцгеймера: алгоритмы машинного обучения на основе маркеров воспаления

Л.В. Андросова<sup>1</sup>, А.Н. Симонов<sup>1</sup>, О.В. Сенько<sup>2</sup>, Н.М. Михайлова<sup>1</sup>, А.В. Кузнецова<sup>3</sup>, Т.П. Ключник<sup>1</sup>

<sup>1</sup>ФГБНУ «Научный центр психического здоровья», Москва, Россия

<sup>2</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, Россия

<sup>3</sup>Институт биохимической физики им. Н.М. Эмануэля Российской академии наук, Москва, Россия

Автор для корреспонденции: Любовь Васильевна Андросова, [androsL@mail.ru](mailto:androsL@mail.ru)

### Резюме

**Обоснование:** болезнь Альцгеймера (БА) как наиболее распространенная форма деменции характеризуется ухудшением познавательных функций и обычно начинается с потери памяти о недавних событиях. Важен поиск биологических методов, чувствительных и доступных, которые можно было бы использовать для ранней диагностики БА и определения тяжести заболевания. **Цель исследования:** разработка алгоритмов машинного обучения (МО) на основе таких воспалительных маркеров, как энзиматическая активность лейкоцитарной эластазы (ЛЭ) и функциональная активность  $\alpha$ 1-протеиназного ингибитора ( $\alpha$ 1-ПИ) для диагностики и оценки тяжести БА. **Пациенты и методы:** в исследование включены 128 человек в возрасте от 55 до 94 лет ( $73,7 \pm 7,9$  года), из которых 91 пациент с диагнозом болезни Альцгеймера и 37 условно здоровых людей (контроль). В качестве классифицирующих признаков для построения моделей рассматривали показатели ЛЭ и  $\alpha$ 1-ПИ в плазме крови. Для построения модели машинного обучения применяли следующие алгоритмы: метод оптимально достоверных разбиений (Optimal Valid Partition, OVP), логистическая регрессия (LR), метод опорных векторов (SVM), случайный лес (RF), градиент бустинга (GB) и метод статистически взвешенных синдромов (MCBC). Был использован программный пакет Data Master Azforus. Прогностическую эффективность построенных классификаторов оценивали по общей точности (accuracy), чувствительности (sensitivity), специфичности (specificity), F-мере и ROC-анализу. **Результаты:** созданные алгоритмы машинного обучения позволили надежно разделить общую группу исследуемых (пациенты + условно здоровые), а также пациентов с различной тяжестью БА на 4 квадранта двумерной диаграммы в координатах ЛЭ и  $\alpha$ 1-ПИ и показали близкую и достаточно высокую прогностическую эффективность. **Заключение:** разработанные алгоритмы машинного обучения оказались высокоэффективными в оценке тяжести БА на основе воспалительных маркеров (энзиматической активности ЛЭ и функциональной активности  $\alpha$ 1-ПИ) и могут быть полезными для ранней диагностики заболевания и своевременного назначения терапии.

**Ключевые слова:** болезнь Альцгеймера, машинное обучение, метод оптимальных достоверных разбиений, алгоритмы бинарной классификации, активность лейкоцитарной эластазы,  $\alpha$ 1-протеиназный ингибитор

**Для цитирования:** Андросова Л.В., Симонов А.Н., Сенько О.В., Михайлова Н.М., Кузнецова А.В., Ключник Т.П. Диагностика и оценка тяжести болезни Альцгеймера: алгоритмы машинного обучения на основе маркеров воспаления. *Психиатрия*. 2024;22(1):6–14. <https://doi.org/10.30629/2618-6667-2024-22-1-6-14>

RESEARCH

UDC 616.89; 571.27

<https://doi.org/10.30629/2618-6667-2024-22-1-6-14>

## Diagnostics and Assessment of the Severity of Alzheimer's Disease: Machine Learning Algorithms Based on Markers of Inflammation

L.V. Androsova<sup>1</sup>, A.N. Simonov<sup>1</sup>, O.V. Senko<sup>2</sup>, N.M. Mikhaylova<sup>1</sup>, A.V. Kuznetsova<sup>3</sup>, T.P. Klyushnik<sup>1</sup>

<sup>1</sup>FSBSI "Mental Health Research Centre", Moscow, Russia

<sup>2</sup>Federal Research Center "Informatics and Management", Russian Academy of Sciences, Moscow, Russia

<sup>3</sup>N.M. Emanuel Institute of Biochemical Physics (IBCP) Organization, Russian Academy of Sciences, Moscow, Russia

Corresponding author: Lubov V. Androsova, [androsL@mail.ru](mailto:androsL@mail.ru)

### Summary

**Background:** as the most common form of dementia, Alzheimer's disease (AD) is characterized by cognitive deterioration and usually begins with loss of memory of recent events. It is important to search for biological, sensitive and affordable methods that could be used for early diagnostics of AD and determine the severity of the disease. **Objective:** to develop machine learning algorithms based on such inflammatory markers as the enzymatic activity of leukocyte elastase (LE) and the functional activity of the  $\alpha$ 1-proteinase inhibitor ( $\alpha$ 1-PI) for diagnosing and assessing the severity of AD. **Patients and methods:** the study included

128 people aged 55 to 94 years ( $73.7 \pm 7.9$  years), of which 91 patients were diagnosed with Alzheimer's disease and 37 apparently healthy people (control). The indicators of LE and  $\alpha 1$ -PI in blood plasma were used as classifying features for building models. The following algorithms were used to build a machine learning model: Optimal Valid Partition (OVP), logistic regression (LR), support vector machine (SVM), random forest (RF), gradient boosting (GB) and statistically weighted syndromes (WSWS). The predictive performance of the constructed classifiers was evaluated by the overall accuracy (accuracy), sensitivity (sensitivity), specificity (specificity), F-measure and ROC-analysis. **Results:** the developed machine learning algorithms made it possible to reliably divide the general group of subjects (patients + conditionally healthy), as well as patients with different AD severity, into 4 quadrants of a two-dimensional diagram in the LE and  $\alpha 1$ -PI coordinates and showed close and fairly high predictive efficiency. **Conclusion:** the developed machine learning algorithms have proven close and sufficiently high prognostic efficacy for assessing the severity of AD based on inflammatory markers (enzymatic activity of LE and functional activity of  $\alpha 1$ -PI) and, probably, can be useful for early diagnostics of the disease and timely administration of therapy.

**Keywords:** Alzheimer's disease, machine learning, optimal reliable partitioning, binary classification algorithms, leukocyte elastase activity,  $\alpha 1$ -proteainase inhibitor

**For citation:** Androsova L.V., Simonov A.N., Senko O.V., Mikhaylova N.M., Kuznetsova A.V., Klyushnik T.P. Diagnostics and Assessment of the Severity of Alzheimer's Disease: Machine Learning Algorithms Based on Markers of Inflammation. *Psychiatry (Moscow) (Psikhiatriya)*. 2024;22(1):6–14. (In Russ.). <https://doi.org/10.30629/2618-6667-2024-22-1-6-14>

## ВВЕДЕНИЕ

Болезнь Альцгеймера (БА) — наиболее распространенная причина деменции в позднем возрасте. Заболевание обычно начинается с ослабления памяти на недавние события и характеризуется неуклонно прогрессирующим ухудшением познавательных функций и нарушениями социальной адаптации [1]. Морфологическим субстратом в тканях головного мозга при данной патологии является повышение уровня амилоида- $\beta$  (А $\beta$ ), составляющего внеклеточные сенильные бляшки, и гиперфосфорилирование тау-белка (p-тау), ассоциированного с микротрубочками. Чрезмерное фосфорилирование тау-белка приводит к его агрегации и переходу в нерастворимую форму с образованием нейрофибриллярных клубков (NFT) [2].

В настоящее время в мире насчитывается около 50 млн человек с деменцией, а к 2050 г. прогнозируется увеличение пациентов этой когорты до 131 млн. Снижение их способности к самообслуживанию ляжет тяжелым бременем на медицинские службы, а также приведет к значительному увеличению материальных затрат [3]. Используемые для лечения БА препараты могут лишь ненадолго замедлить развитие этого заболевания и ослабить прогрессирование на начальных этапах деменции [4, 5]. Тяжесть деменции зависит от этапа заболевания и скорости прогрессирования. Предполагается, что лечение заболевания на самых ранних этапах может оказать временный эффект. В связи с этим чрезвычайно актуален поиск чувствительных и доступных методов, которые могут быть использованы для ранней доклинической диагностики БА и/или оценки тяжести уже развившегося заболевания.

Общепризнанным результатом исследований последних десятилетий стало представление о том, что важным патогенетическим звеном БА является нейровоспаление, ассоциированное с повышением воспалительных маркеров в крови пациентов [6–8].

В наших предыдущих исследованиях было показано повышение уровня медиаторных и острофазных молекул (ИЛ-6, СРБ и  $\alpha 1$ -ПИ) в крови пациентов, а также

снижение протеолитической активности ЛЭ по сравнению с контролем, коррелирующие со степенью снижения когнитивных функций [9]. ЛЭ — сериновая протеаза, содержится в азурофильных гранулах нейтрофилов, секретируется во внеклеточное пространство при активации этих клеток и считается одним из главных участников воспаления. Основным регулятором активности данного фермента является  $\alpha 1$ -протеиназный ингибитор ( $\alpha 1$ -ПИ), который способствует ограничению очага воспаления. Снижение активности ЛЭ при БА, предположительно, отражает нарушение проницаемости гематоэнцефалического барьера (ГЭБ), что способствует трансмиграции нейтрофилов в мозг [10–13].

В настоящем исследовании предпринята попытка создания алгоритмов машинного обучения на основе воспалительных маркеров ЛЭ и  $\alpha 1$ -ПИ для диагностики тяжести БА. При этом необходимо иметь в виду, что существует большое количество алгоритмов бинарной классификации машинного обучения [14] и неизвестно, какой из них окажется наилучшим для решения конкретной задачи. Согласно NFL-теореме [15] универсальных алгоритмов, решающих разнообразные задачи с одинаково высокой точностью, не существует. Поэтому в данной работе использовали несколько алгоритмов бинарной классификации машинного обучения: метод оптимальных достоверных разбиений (ОДР); логистическую регрессию; метод опорных векторов (Support Vector Machine, SVM); случайный решающий лес (RF); градиентный бустинг (GB), генерирующий решающий лес, и метод статистически взвешенных синдромов (МСВС) [16, 17].

Для оценки прогностической эффективности алгоритмов бинарной классификации использовали несколько видов метрик: общая точность (accuracy), чувствительность (sensitivity), специфичность (specificity), F-меру и ROC-анализ.

**Целью** настоящей работы явилась разработка модели машинного обучения при помощи алгоритмов бинарной классификации для отражения тяжести БА с использованием таких маркеров воспаления, как ЛЭ и  $\alpha 1$ -ПИ.

## КЛИНИЧЕСКАЯ ХАРАКТЕРИСТИКА ПАЦИЕНТОВ И МЕТОДЫ ИССЛЕДОВАНИЯ

Исследование проводилось в лаборатории нейробиологии (заведующий лабораторией профессор Т.П. Ключник) и лаборатории доказательной медицины и биостатистики (заведующий лабораторией А.Н. Симонов). Пациенты с БА были обследованы и проходили амбулаторное или стационарное лечение в отделе гериатрической психиатрии (заведующий отделом профессор С.И. Гаврилова) ФГБНУ «Научный центр психического здоровья» (директор профессор Т.П. Ключник).

**Критерии включения:** возраст на момент обследования 55 лет и старше; диагноз деменции при болезни Альцгеймера.

**Критерии не включения:** клинические и/или лабораторные признаки инфекционной или аутоиммунной патологии в течение двух месяцев до обследования.

### Характеристика клинического материала

Группу пациентов с болезнью Альцгеймера составил 91 человек (56 мужчин и 35 женщин) в возрасте от 55 до 94 лет ( $73,7 \pm 7,9$  года). Диагноз болезни Альцгеймера был установлен согласно критериям ICD-10 и NINCDS-ADRDA [18]. Степень тяжести деменции — мягкую ( $n = 23$ ), умеренную ( $n = 43$ ) и тяжелую ( $n = 25$ ) — определяли по клиническим признакам с использованием шкалы CDR (Clinical Dementia Rating) [19], а также суммарного балла мини-теста психического состояния (Mini Mental State Examination, MMSE) [20]. Суммарный балл по шкале MMSE был в пределах 16–28 баллов при мягкой деменции, от 16 до 24 баллов — при умеренной и от 0 до 15 баллов при тяжелой. При разной тяжести деменции показатели оценочных шкал могут частично перекрываться.

Контрольную группу составили 37 практически здоровых лиц (18 мужчин и 19 женщин) в возрасте от 55 до 79 лет ( $64,6 \pm 6,6$  года), при диспансерном обследовании не обнаруживших признаков психических расстройств, обострений соматических заболеваний или симптомов острой инфекции.

### Этические аспекты

Все обследуемые давали информированное письменное согласие на участие в исследовании. Проведение исследования соответствовало положениям Хельсинкской декларации 1964 г., пересмотренной в 1975–2013 гг., и одобрено Локальным этическим комитетом ФГБНУ НЦПЗ (протокол № 194 от 23.03.2014).

### Ethic aspects

All examined participants of study signed the informed consent to take part in a study. This study complies with the Principles of the WMA Helsinki Declaration 1964 amended 1975–2013. The research protocol was approved by Local Ethical Committee of FSBSI “Mental Health Research Centre” (protocol # 194 from March 23, 2014).

### Иммунологическое исследование

Уровень воспалительных маркеров определяли в плазме крови, которую получали стандартным методом, используя вакутейнеры с напылением ЭДТА (КЗЕ, К<sub>3</sub>EDTA).

**Энзиматическую активность ЛЭ** определяли ферментативным спектрофотометрическим методом с использованием специфического субстрата N-терт-бутоксикарбонил-аланин-β-нитрофенилового эфира (BOC-Ala-ONp) и оценивали в нмоль/мин × мл (чувствительность метода 40 нмоль/мин × мл) [21].

**Функциональную активность α1-ПИ** определяли спектрофотометрическим методом и оценивали в ИЕ/мл (ингибиторные единицы/мл) (чувствительность метода 5 ИЕ/мл) [22].

### Алгоритмы машинного обучения

Для построения прогностических моделей использовались подходы, основанные на методе оптимально достоверных разбиений (ОДР), логистической регрессии, методе опорных векторов (SVM), случайном решающем лесе, градиентном бустинге, генерирующий решающий лес, и методе статистически взвешенных синдрома (МСВС).

**Метод оптимальных достоверных разбиений (ОДР).** Метод позволяет провести разбиение исследуемой выборки (пациенты + здоровые) на подгруппы, близкие по клинико-иммунологическим показателям. Верификация закономерностей, характеризуемых оптимальными разбиениями, производится с помощью специального варианта перестановочного теста [17]. ОДР можно рассматривать как звено между статистическими методами и машинным обучением. Он как бы разлагает сложную модель на множество простых оптимальных статистически значимых моделей без избыточных элементов, вклад каждого элемента модели проверяется на статистическую значимость.

### Сравнительная оценка прогностической эффективности полученных моделей на тестовой выборке

Оценка прогностической эффективности полученных моделей дает возможность судить об их практической пригодности. Для оценки и сравнения прогностической эффективности построенных классификаторов использовали несколько видов метрик: точность, чувствительность, специфичность, F-меру и ROC-анализ.

Результатом работы любых классификаторов являются следующие четыре основных характеристики (числа), на основе которых строят различные классификационные метрики.

1. TP (True Positive) истинно положительный исход — количество правильно классифицированных пациентов.
2. TN (True Negative) истинный отрицательный исход — количество ложно классифицированных пациентов.
3. FP (False Positive) ложноположительный исход — количество неверно классифицированных пациентов как здоровых.

**Таблица 1.** Сопряженность или матрица ошибок  
**Table 1** Conjugacy or error matrix

Предсказание/Prediction	Наблюдение/Observation		Сумма/Sum
	есть заболевание/there is a disease	нет заболевания/no disease	
есть заболевание/there is a disease	True Positive (TP)	False Positive (FP)	TP + FP
нет заболевания/no disease	False Negative (FN)	True Negative (TN)	FN + TN
сумма/sum	TP + FN	TN + FP	
	Sensitivity TPR = TP/(TP + FN)	Specificity TNR = TN/(TN + FP)	

4. FN (False Negative) ложноотрицательный исход — количество неправильно классифицированных здоровых как больных.

Эти четыре числа представлены в виде таблицы (табл. 1) сопряженности или матрицы ошибок (confusion matrix).

На основании этих четырех чисел строятся следующие метрики:

- чувствительность (true positive rate (TPR)) — число правильных положительных прогнозов, деленное на общее количество положительных исходов, наилучшая чувствительность равна 1,0, а наихудшая — 0,0.  $TPR = TP / (TP + FN)$ ;
- специфичность (true negative rate (TNR)) — количество правильных отрицательных прогнозов, деленное на общее количество отрицательных исходов, лучшая специфичность — 1,0, а худшая — 0,0.  $TNR = TN / (TN + FP)$ ;
- общая точность (ассигасу) рассчитывается как число всех правильных прогнозов, деленное на общее количество исходов, иначе, отношение правильно классифицированных пациентов к общему количеству пациентов,  $ассигасу = (TP + TN) / (TP + FP + FN + TN)$ , наилучшая точность равна 1,0, а наихудшая — 0,0;
- F-мера вычисляется по следующей формуле:  $F = 2TP / (2TP + FP + FN)$ .

#### ROC-анализ

Универсальным способом сравнения точности алгоритмов считается ROC-анализ. Кривая ROC — это способ визуализации эффективности бинарных классификатора. Количественным описанием ROC-кривой является площадь под этой кривой (AUC). Величина AUC, равная 1, соответствует идеальному классификатору, а значение, близкое к 0,5, признается плохой моделью, поскольку это эквивалентно случайной классификации. Принята следующая система оценки эффективности классификаторов по величине AUC: 0,90–1 — отличный классификатор; 0,80–0,90 — хороший классификатор; 0,70–0,80 — удовлетворительный классификатор; 0,60–0,70 — плохой классификатор; 0,50–0,60 — неудачный классификатор.

**Статистическую обработку** данных проводили с использованием программ R (R version 3.6.1) и STATA (version 12.1). Уровень статистической значимости был выбран при  $p < 0,05$ .

Графическое изображение распределения исследуемых (пациенты + контроль) представлено в виде

двумерной диаграммы со значениями ЛЭ (горизонтальная ось) и  $\alpha 1$ -ПИ (вертикальная ось).

Значимости распределений признаков ЛЭ и  $\alpha 1$ -ПИ по четырем квадрантам оцениваются с помощью метода Краскела–Уоллиса с поправками на множественные сравнения по методу Sidak [23].

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### Анализ методом ОДР

Метод оптимальных достоверных разбиений (ОДР) позволил выделить на двумерных диаграммах рассеяния области, где преобладают пациенты с близкими иммунологическими характеристиками.

На рис. 1 представлена двумерная диаграмма рассеяния изучаемой выборки в координатах ЛЭ (горизонтальная ось) и  $\alpha 1$ -ПИ (вертикальная ось).

Как видно из рис. 1, все обследованные (пациенты + контроль) по иммунологическим показателям распределились неравномерно по 4 квадрантам. Пациенты с БА распределились в основном в 1-м (ЛЭ < 181,2 и  $\alpha 1$ -ПИ > 45,05), 2-м (ЛЭ > 181,2 и  $\alpha 1$ -ПИ > 45,05) и 4-м (ЛЭ < 181,2 и  $\alpha 1$ -ПИ < 45,05) квадрантах, что свидетельствует о гетерогенности группы пациентов.

Группа контроля сосредоточилась преимущественно (94,6%) в правом нижнем квадранте (3-й) (ЛЭ > 181,2 и  $\alpha 1$ -ПИ < 45,05). В этот квадрант попали также 13 (14,3%) пациентов с болезнью Альцгеймера.

В табл. 2 приведены характеристики иммунологических показателей квадрантов, полученные методом машинного обучения.

Как видно из табл. 2, 1-й квадрант характеризуется повышенной активностью  $\alpha 1$ -ПИ, отражающей высокий уровень воспаления, и низкой активностью ЛЭ, выходящей за пределы нижнего диапазона значений контрольной группы (условно здоровых людей). В этот квадрант вошли 27 пациентов с БА (29,7% общего числа больных), из них 24 с умеренной и тяжелой стадией заболевания (88,9%) и три с легкой стадией (11,1%). Показатели этой группы пациентов характеризуются высоким уровнем воспаления и, предположительно, критическим нарушением проницаемости ГЭБ, в качестве маркера которого выступает низкий уровень ЛЭ.

2-й квадрант также отличается высокой активностью  $\alpha 1$ -ПИ и равной/повышенной по сравнению с контролем активностью ЛЭ. В этот квадрант вошли 38 пациентов с БА, из них 31 — с легкой и умеренной степенью деменции (81,6% от абсолютного числа

**Таблица 2.** Иммунологические характеристики квадрантов (медиана [25-й; 75-й перцентиль]), минимум-максимум  
**Table 2** Immunological characteristics of the quadrants (median [25; 75 percentile]), minimum-maximum

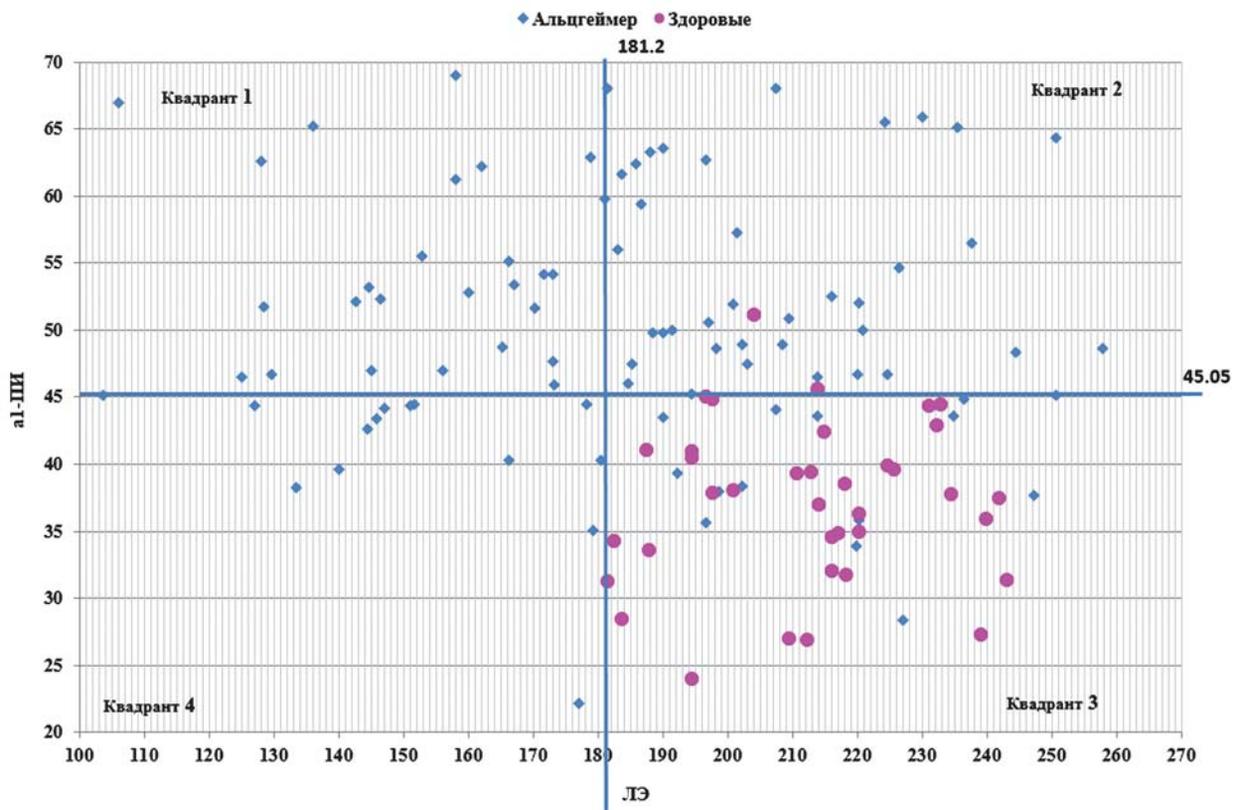
Показатели/Indicators	1-й квадрант/ 1 <sup>st</sup> quadrant (n = 27)	2-й квадрант/ 2 <sup>nd</sup> quadrant (n = 40)	3-й квадрант/ 3 <sup>rd</sup> quadrant (n = 48)	4-й квадрант/ 4 <sup>th</sup> quadrant (n = 13)
Активность ЛЭ, нмоль/мин × мл/ Activity of LE, nmol/min × ml	158 [136,1; 170,2] 103,7–181,0 P (1–2) = 0,00000 P (1–3) = 0,00000 P (1–4) = 0,732334	203,6 [190; 222,6] 181,4–257,9 P (2–3) = 0,214506 P (2–4) = 0,000000	214,4 [197,1; 226,4] 181,4–247,3 P (3–4) = 0,000000	151 [144,4; 177] 127,0–180,4
Функциональная активность α1-ПИ, ИЕ/мл/ Functional activity α1-PI, UI/ml	53,2 [47,7; 61,2] 45,1–69,0 P (1–2) = 0,689200 P (1–3) = 0,00000 P (1–4) = 0,00000	51,05 [48,4; 62,0] 45,1–68,0 P (2–3) = 0,00000 P (2–4) = 0,00000	37,85 [34,07; 41,0] 24–45 P (3–4) = 0,023461	42,6 [39,6; 44,4] 22,2–44,5

больных), 7 пациентов — с тяжелой стадией заболевания (18,4%) и 2 человека из контрольной группы (5,4%). Эту группу отличает высокий уровень воспаления и относительная сохранность ГЭБ.

3-й квадрант включает значения активности α1-ПИ и ЛЭ, находящиеся в пределах контрольного диапазона. В этот квадрант вошли показатели 35 человек из контрольной группы (94, 6%) и 13 пациентов с БА, из них 8 — с мягкой деменцией (61,5%), 4 — с умеренной (30,8%) и 1 — с тяжелой деменцией (7,7%). В этой группе отсутствуют отклонения от нормы исследуемых признаков воспаления.

4-й квадрант отличает умеренное повышение активности α1-ПИ и низкая активность ЛЭ, выходящая за пределы нижнего диапазона контрольных значений. В этот квадрант вошли показатели 13 пациентов с БА (14,3%), среди них 12 пациентов с умеренной и тяжелой стадией заболевания (92,3%) и 1 пациент с мягкой деменцией (7,7%). Их характеризует умеренный уровень воспаления и, предположительно, критическое нарушение проницаемости ГЭБ, аналогично показателям пациентам из 1-го квадранта.

Проанализировано также изменение изучаемых показателей когнитивного функционирования пациентов



**Рис. 1.** Двумерная диаграмма рассеяния изучаемой выборки в координатах ЛЭ (горизонтальная ось) и α1-ПИ (вертикальная ось)

**Fig. 1** Two-dimensional scatterplot of the studied sample in the coordinates LE (horizontal axis) and α1-PI (vertical axis)

**Таблица 3.** Уровень когнитивных нарушений по суммарному баллу шкалы MMSE у пациентов по квадрантам (медиана [25-й; 75-й перцентиль]), минимум-максимум

**Table 3** The level of cognitive impairment according to MMSE scores in patients by quadrants (median [25; 75 percentile]), minimum-maximum

Показатели/Indicators	1-й квадрант/ 1 <sup>st</sup> quadrant (n = 27)	2-й квадрант/ 2 <sup>nd</sup> quadrant (n = 38)	3-й квадрант/ 3 <sup>rd</sup> quadrant (n = 13)	4-й квадрант/ 4 <sup>th</sup> quadrant (n = 13)
Мини-тест психического состояния/MMSE	14,0 [9; 17] 1–28 P (1–2) = 0,049099 P (1–3) = 0,006690 P (1–4) = 0,886726	17 [12; 22] 0–27 P (2–3) = 0,362769 P (2–4) = 0,118001	19 [17; 22] 6–24 P (3–4) = 0,014044	14 [10; 17] 0–23

**Таблица 4.** Эффективность работы классификаторов

**Table 4** The efficiency of the classifiers

Метод МО/Machine learning method	Точность/ Accuracy	Чувствительность/ Recall	Специфичность/ Specificity	F-мера/ F-measure	ROC AUC
Логистическая регрессия/Logistics regression	0,828	0,65	0,90	0,66	0,899
Метод опорных векторов/SVM	0,844	0,73	0,89	0,73	0,89
Решающий лес/Decision forest	0,836	0,78	0,86	0,73	0,89
Градиентный бустинг/Gradient boosting	0,828	0,76	0,86	0,72	0,88
Статистически взвешенные синдромы (CBC)/ Statistically weighted syndromes (SWS)	0,836	0,84	0,84	0,75	0,88

в выделенных квадрантах (по баллам шкалы MMSE) (табл. 3).

Как видно из табл. 3, количества баллов по шкале MMSE самые минимальные в 1-м и 4-м квадранте, что связано с более тяжелой стадией заболевания пациентов с БА в отличие от суммы баллов MMSE у пациентов, чьи показатели оказались во 2-м и в 3-м квадрантах.

Использование пяти различных методов машинного обучения, включая логистическую регрессию, метод опорных векторов (SVM), случайный решающий лес (random forest), градиентный бустинг, генерирующий решающий лес (decision forest), статистически взвешенные синдромы, показало близкую и достаточно высокую прогностическую эффективность. Общая точность — 0,828–0,844, чувствительность — 0,65 ÷ 0,84, специфичность — 0,84 ÷ 0,9, F — 0,66 ÷ 0,75 и ROC AUC — 0,88 ÷ 0,9 (табл. 3).

В табл. 4 приведены результаты анализа пяти алгоритмов машинного обучения, подтверждающие их высокую прогностическую эффективность и сходство.

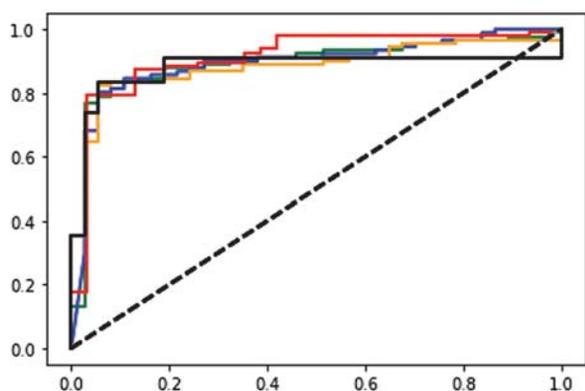
Графически качество всех пяти классификаторов представлено ROC-кривыми на рис. 2.

Таким образом, метод оптимально достоверных разбиений с использованием воспалительных маркеров еще раз подтвердил гетерогенность группы пациентов с БА по иммунологическим показателям. В целом распределение показателей пациентов по выделенным квадрантам соответствовало различной степени тяжести БА, что свидетельствует о том, что активность ЛЭ и α1-ПИ являются адекватными и высокоэффективными классифицирующими признаками для оценки тяжести БА.

Сравнение полученных в данном исследовании диагностических алгоритмов с алгоритмами, разработанными на основе показателей ПЭТ (позитронно-эмиссионная томография) методом глубокой сверточной нейронной сети (CNN), приведенными в работе Н. Choi и К.Н. Jin (2018) [24], показало, что алгоритмы на основе выбранных иммунологических показателей значительно превосходят алгоритмы с использованием показателей ПЭТ по трудозатратам и общей стоимости диагностики.

### ЗАКЛЮЧЕНИЕ

Разработанные алгоритмы машинного обучения показали близкую и достаточно высокую прогностическую эффективность оценки тяжести БА на основе воспалительных маркеров (энзиматической активности ЛЭ и функциональной активности α1-ПИ) и могут быть полезными для ранней диагностики заболевания и своевременного назначения терапии. Дальнейшее



**Рис. 2.** ROC-кривые для пяти классификаторов

**Fig. 2** ROC curves for five classifiers

развитие этого исследования предполагает применение разработанных алгоритмов для пациентов с мягким когнитивным снижением.

### СПИСОК ИСТОЧНИКОВ/REFERENCES

- Zhang XX, Tian Y, Wang ZT, Ma YH, Tan L, Yu JT. The Epidemiology of Alzheimer's Disease Modifiable Risk Factors and Prevention. *J Prev Alzheimers Dis.* 2021;8(3):313–321. doi: [10.14283/jpad.2021.15](https://doi.org/10.14283/jpad.2021.15) PMID: 34101789.
- Winblad B, Amouyel P, Andrieu S, Ballard C, Brayne C, Brodaty H, Cedazo-Minguez A, Dubois B, Edvardsson D, Feldman H, Fratiglioni L, Frisoni GB, Gauthier S, Georges J, Graff C, Iqbal K, Jessen F, Johansson G, Jönsson L, Kivipelto M, Knapp M, Mangialasche F, Melis R, Nordberg A, Rikkert MO, Qiu C, Sakmar TP, Scheltens P, Schneider LS, Sperling R, Tjernberg LO, Waldemar G, Wimo A, Zetterberg H. Defeating Alzheimer's disease and other dementias: a priority for European science and society. *The Lancet Neurology.* 2016;15(5):455–532. doi: [10.1016/S1474-4422\(16\)00062-4](https://doi.org/10.1016/S1474-4422(16)00062-4) PMID: 26987701.
- Karagiannidou M, Wittenberg R, Landeiro FIT, Park AL, Fry A, Knapp M, Gray AM, Tockhorn-Heidenreich A, Castro Sanchez AY, Ghinai I, Handels R, Lecomte P, Wolstenholme J; ROADMAP Group. Systematic literature review of methodologies and data sources of existing economic models across the full spectrum of Alzheimer's disease and dementia from apparently healthy through disease progression to end of life care: a systematic review protocol. *BMJ Open.* 2018;8(6):e020638. doi: [10.1136/bmjopen-2017-020638](https://doi.org/10.1136/bmjopen-2017-020638) PMID: 29884696; PMCID: PMC6009454.
- Breijyeh Z, Karaman R. Comprehensive Review on Alzheimer's Disease: Causes and Treatment. *Molecules.* 2020;25(24):5789. doi: [10.3390/molecules25245789](https://doi.org/10.3390/molecules25245789) PMID: 33302541; PMCID: PMC7764106.
- Passeri E, Elkhoury K, Morsink M, Broersen K, Linder M, Tamayol A, Malaplate C, Yen FT, Arab-Tehrany E. Alzheimer's Disease: Treatment Strategies and Their Limitations. *Int J Mol Sci.* 2022;23(22):13954. doi: [10.3390/ijms232213954](https://doi.org/10.3390/ijms232213954) PMID: 36430432; PMCID: PMC9697769.
- Bettcher BM, Johnson SC, Fitch R, Casaletto KB, Heffernan KS, Asthana S, Zetterberg H, Blennow K, Carlsson CM, Neuhaus J, Bendlin BB, Kramer JH. Cerebrospinal Fluid and Plasma Levels of Inflammation Differentially Relate to CNS Markers of Alzheimer's Disease Pathology and Neuronal Damage. *J Alzheimers Dis.* 2018;62(1):385–397. doi: [10.3233/JAD-170602](https://doi.org/10.3233/JAD-170602) PMID: 29439331; PMCID: PMC6007886.
- Popp J, Oikonomidi A, Tautvydaitė D, Dayon L,bacher M, Migliavacca E, Henry H, Kirkland R, Severin I, Wojcik J, Bowman GL. Markers of neuroinflammation associated with Alzheimer's disease pathology in older adults. *Brain Behav Immun.* 2017;62:203–211. doi: [10.1016/j.bbi.2017.01.020](https://doi.org/10.1016/j.bbi.2017.01.020) Epub 2017 Feb 1. PMID: 28161476.
- Schram MT, Euser SM, de Craen AJ, Witteman JC, Frölich M, Hofman A, Jolles J, Breteler MM, Westendorp RG. Systemic markers of inflammation and cognitive decline in old age. *J Am Geriatr Soc.* 2007;55(5):708–716. doi: [10.1111/j.1532-541.2007.01159.x](https://doi.org/10.1111/j.1532-541.2007.01159.x)
- Андросова ЛВ, Михайлова НМ, Зозуля СА, Дупин АМ, Ключник ТП. Иммунобиохимические маркеры воспаления при деменциях, ассоциированных с возрастом. *Российский психиатрический журнал.* 2017;(4):61–66. Андросова LV, Mikhaylova NM, Zozulya SA, Dupin AM, Klyushnik TP. Immunobiochemical markers of inflammation in development of age-associated dementia. *Russian Journal of Psychiatry.* 2017;(4):61–66. (In Russ.).
- Shimamura A, Kamanaka Y, Ikeda Y, Kondo K, Suzuki Y, Umemura K. Neutrophil elastase inhibition reduces cerebral ischemic damage in the middle cerebral artery occlusion. *Brain Res.* 2000;858(1):55–60. doi: [10.1016/s0006-8993\(99\)02431-2](https://doi.org/10.1016/s0006-8993(99)02431-2) PMID: 10700596.
- Ushakumari CJ, Zhou QL, Wang YH, Na S, Rigor MC, Zhou CY, Kroll MK, Lin BD, Jiang ZY. Neutrophil Elastase increases vascular permeability and leukocyte transmigration in cultured endothelial cells and obese mice. *Cells.* 2022;11(15):2288. doi: [10.3390/cells11152288](https://doi.org/10.3390/cells11152288) PMID: 35892585; PMCID: PMC9332277.
- Smyth LCD, Murray HC, Hill M, van Leeuwen E, Highet B, Magon NJ, Osanlouy M, Mathiesen SN, Mockett B, Singh-Bains MK, Morris VK, Clarkson AN, Curtis MA, Abraham WC, Hughes SM, Faull RLM, Kettle AJ, Dragunow M, Hampton MB. Neutrophil-vascular interactions drive myeloperoxidase accumulation in the brain in Alzheimer's disease. *Acta Neuropathol Commun.* 2022;10(1):38. doi: [10.1186/s40478-022-01347-2](https://doi.org/10.1186/s40478-022-01347-2) PMID: 35331340; PMCID: PMC8944147.
- Stalder AK, Ermini F, Bondolfi L, Krenger W, Burbach GJ, Deller T, Coomaraswamy J, Staufenbiel M, Landmann R, Jucker M. Invasion of hematopoietic cells into the brain of amyloid precursor protein transgenic mice. *J Neurosci.* 2005;25(48):11125–11132. doi: [10.1523/JNEUROSCI.2545-05.2005](https://doi.org/10.1523/JNEUROSCI.2545-05.2005) PMID: 16319312; PMCID: PMC6725647.
- Fernandez-Delgado M, Cernadas E, Barro S. Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research.* 2014;15:3133–3181. Submitted 11/13; Revised 4/14; Published 10/14.
- Wolpert DH, Macready WG. No Free Lunch Theorems for Optimization. *IEEE Transactions on Evolutionary Computation.* 1997;1:67.
- Kuznetsova AV, Kostomarova IV, Senko OV. Modification of the method of optimal valid partitioning for comparison of patterns related to the occurrence of ischemic stroke in two groups of patients. *Pattern*

- Recognition and Image Analysis*. 2014;24(1):114–123. doi: [10.1134/S105466181401009X](https://doi.org/10.1134/S105466181401009X)
17. Senko OV, Kuznetsova AV. A recognition method based on collective decision making using systems of regularities of various types. *Pattern Recognition and Image Analysis*. 2010;20(2):152–162. doi: [10.1134/S1054661810020069](https://doi.org/10.1134/S1054661810020069)
  18. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984;34(7):939–944. doi: [10.1212/wnl.34.7.93](https://doi.org/10.1212/wnl.34.7.93). PMID: 6610841.
  19. Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. *Br J Psychiatr*. 1982;140:3:566–572. doi: [10.1192/bjp.140.6.566](https://doi.org/10.1192/bjp.140.6.566) PMID: 7104545.
  20. Folstein MF, Folstein SF, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiat Res*. 1975;12:189–198.
  21. Доценко ВЛ, Нешкова ЕА, Яровая ГА. Выявление лейкоцитарной эластазы человека из комплекса с плазменным  $\alpha$ 1-протеиназным ингибитором по ее энзиматической активности с синтетическим субстратом. *Вопр. мед. хим.* 1994;40(3):20–25. Dotsenko VL, Neshkova EA, Jarovaia GA. Docenko VL, Neshkova EA, Jarovaja GA. Vyjavlenie lejkocitarno jelastazy cheloveka iz kompleksa s plazmennym  $\alpha$ 1-proteinaznym ingibitorom po ee jenzimaticheskoj aktivnosti s sinteticheskim substratom. [Detection of human leukocyte elastase from a plasma alpha-1-proteinase inhibitor complex by its enzymatic activity with a synthetic substrate]. *Vopr Med Khim*. 1994;40(3):20–25. (In Russ.). PMID: 8079434.
  22. Нартикова ВФ, Пасхина ТС. Унифицированный метод определения активности  $\alpha$ 1-антитрипсина и  $\alpha$ 2-макроглобулина активности в сыворотке крови человека (плазмы). *Вопр. мед. хим.* 1979;25(4):494–499. Nartikova VF, Pashkina TS. Unifitsirovannyi metod opredeleniia aktivnosti al'fa 1-antitripsina i al'fa 2-makroglobulina v syvorotke (plazme) krovi cheloveka [Uniform method for determining the alpha 1-antitrypsin and alpha 2-macroglobulin activity in human blood serum (plasma)]. *Vopr Med Khim*. 1979;25(4):494–499. (In Russ.). PMID: 89758.
  23. Sidák ZK. Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. *Journal of the American Statistical Association*. 1967;62(318):626–633. doi: [10.1080/01621459.1967.10482935](https://doi.org/10.1080/01621459.1967.10482935)
  24. Choi H, Jin KH. Alzheimer's Disease Neuroimaging Initiative. Predicting cognitive decline with deep learning of brain metabolism and amyloid imaging. *Behav Brain Res*. 2018;344:103–109. doi: [10.1016/j.bbr.2018.02.017](https://doi.org/10.1016/j.bbr.2018.02.017) Epub 2018 Feb 14. PMID: 29454006.

#### Сведения об авторах

Любовь Васильевна Андросова, кандидат биологических наук, ведущий научный сотрудник, лаборатории нейроиммунологии, ФГБНУ «Научный центр психического здоровья», Москва, Россия, <https://orcid.org/0000-0002-2433-8810>

[androsL@mail.ru](mailto:androsL@mail.ru)

Анатолий Никифорович Симонов, кандидат биологических наук, заведующий лабораторией, лаборатория доказательной медицины и биостатистики, ФГБНУ «Научный центр психического здоровья», Москва, Россия, <https://orcid.org/0000-0003-0564-932X>

[simonov1951@rambler.ru](mailto:simonov1951@rambler.ru)

Олег Валентинович Сенько, доктор физико-математических наук, ведущий научный сотрудник, Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, Россия, <https://orcid.org/0000-0002-5586-3503>

[senkoov@mail.ru](mailto:senkoov@mail.ru)

Наталья Михайловна Михайлова, доктор медицинских наук, ведущий научный сотрудник, ФГБНУ «Научный центр психического здоровья», Москва, Россия, <https://orcid.org/0000-0003-0184-016X>

[MikhaylovaNM@yandex.ru](mailto:MikhaylovaNM@yandex.ru)

Анна Викторовна Кузнецова, кандидат биологических наук, старший научный сотрудник, лаборатория математической биофизики, ИБХФ им. Н.М. Эмануэля РАН, Москва, Россия, <https://orcid.org/0000-0002-0297-7013>

[azforus@yandex.ru](mailto:azforus@yandex.ru)

Татьяна Павловна Клушник, доктор медицинских наук, профессор, заведующий лабораторией, лаборатория нейроиммунологии, директор, ФГБНУ «Научный центр психического здоровья», Москва, Россия, <https://orcid.org/0000-0001-5148-3864>

[klushnik2004@mail.ru](mailto:klushnik2004@mail.ru)

**Information about the authors**

*Lubov V. Androsova*, Cand. of Sci. (Biol.), Leading Researcher, Laboratory of Neuroimmunology, FSBSI "Mental Health Research Centre", Moscow, Russia, <https://orcid.org/0000-0002-2433-8810>

[androsL@mail.ru](mailto:androsL@mail.ru)

*Anatoly N. Simonov*, Cand. of Sci. (Biol.), Head of the Laboratory, Laboratory of Evidence-Based Medicine and Biostatistics, FSBSI "Mental Health Research Centre", Moscow, Russia, <https://orcid.org/0000-0003-0564-932X>

[simonov1951@rambler.ru](mailto:simonov1951@rambler.ru)

*Oleg V. Senko*, Dr. of Sci. (Physics and Math.), Leading Researcher, Federal Research Center "Informatics and Management", Russian Academy of Sciences, Moscow, Russia, <https://orcid.org/0000-0002-5586-3503>

[senkoov@mail.ru](mailto:senkoov@mail.ru)

*Nataliya M. Mikhaylova*, Dr. of Sci. (Med.), Leading Researcher, FSBSI "Mental Health Research Centre", Moscow, Russia, <https://orcid.org/0000-0003-0184-016X>

[MikhaylovaNM@yandex.ru](mailto:MikhaylovaNM@yandex.ru)

*Anna V. Kuznetsova*, Dr. of Sci. (Biol.), Senior Researcher, Laboratory of Mathematical Biophysics, N.M. Emanuel IBCP of RAS, Moscow, Russia, <https://orcid.org/0000-0002-0297-7013>

[azforus@yandex.ru](mailto:azforus@yandex.ru)

*Tatyana P. Klyushnik*, Dr. of Sci. (Med.), Professor, Head of the Laboratory, Laboratory of Neuroimmunology, Director, FSBSI "Mental Health Research Centre", Moscow, Russia, <https://orcid.org/0000-0001-5148-3864>

[klushnik2004@mail.ru](mailto:klushnik2004@mail.ru)

*Авторы заявляют об отсутствии конфликта интересов.*

*The authors declare no conflict of interests.*

Дата поступления 25.09.2023

Received 25.09.2023

Дата рецензирования 30.10.2023

Revised 30.10.2023

Дата принятия 19.12.2023

Accepted for publication 19.12.2023