



Качан Т.В. ✉, Семенова И.И., Марченко Л.Н., Федулов А.С.

Белорусский государственный медицинский университет, Минск, Беларусь

Клинические и структурные предикторы остроты зрения и фенотипирование возрастной макулярной дегенерации на основе методов искусственного интеллекта

Конфликт интересов: не заявлен.

Вклад авторов: концепция и дизайн исследования, редактирование, обработка данных, работа с базой данных (предобработка, инженерия признаков), построение и валидация моделей машинного обучения, написание текста статьи – Качан Т.В.; концепция и дизайн исследования, сбор клинического материала, редактирование, формирование базы данных – Семенова И.И.; редактирование рукописи, утверждение окончательного варианта для публикации – Марченко Л.Н., Федулов А.С.

Подана: 23.03.2026

Принята: 11.05.2026

Контакты: tvk35@Yahoo.com

Резюме

Цель. Комплексная оценка клинических и структурных предикторов остроты зрения у пациентов с возрастной макулярной дегенерацией (ВМД), построение прогностической модели риска низкого зрения и выделение подтипов заболевания с использованием методов машинного обучения.

Материалы и методы. Ретроспективный анализ данных 463 глаз (278 пациентов) с ВМД. Оценивали демографические, функциональные (максимальная скорректированная острота зрения (МКОЗ)) и структурные параметры оптической когерентной томографии. Применялись: критерий Краскела – Уоллиса для сравнения групп, алгоритм случайного леса (Random Forest) для ранжирования предикторов, логистическая регрессия для прогнозирования риска низкого зрения (МКОЗ <0,1) и метод К-средних для кластеризации фенотипов.

Результаты. Наибольшую значимость в предсказании МКОЗ имели показали: стадия AREDS (Age-Related Eye Disease Study, 0,164), состояние пигментного эпителия (RPE status; 0,110) и центральная толщина сетчатки (ЦТС foveola; 0,106). Модель логистической регрессии продемонстрировала хорошее качество прогноза низкого зрения (AUC=0,81); ключевыми факторами риска явились увеличение среднего объема сетчатки (ОШ=2,24) и наличие фиброваскулярной отслойки (ОШ=1,98). Кластерный анализ выделил три клинических фенотипа: активная неоваскуляризация (умеренное снижение зрения, высокая частота экссудативных изменений), отечный/фиброзный (наиболее тяжелый, с максимальной толщиной сетчатки и худшими функциональными исходами) и дружный/ранний (относительно сохранное зрение, высокая частота друз).

Заключение. Применение методов машинного обучения позволило подтвердить ключевую роль структурной целостности нейросенсорной сетчатки и выделить три клинически значимых фенотипа ВМД. Разработанная прогностическая модель



и фенотипирование могут служить основой для персонализации наблюдения и лечения пациентов.

Ключевые слова: возрастная макулярная дегенерация, машинное обучение, случайный лес, кластеризация, острота зрения, прогнозирование, фенотипы, оптическая когерентная томография

Kachan T. ✉, Semenava I., Marchenko L., Fedulov A.
Belarusian State Medical University, Minsk, Belarus

Clinical and Structural Predictors of Visual Acuity and Phenotyping of Age-Related Macular Degeneration Based on Artificial Intelligence Methods

Conflict of interest: nothing to declare.

Authors' contribution: study concept and design, editing, data processing, database management (preprocessing, feature engineering), construction and validation of machine learning models, and manuscript writing – Kachan T.; study concept and design, clinical data collection, editing, and database development – Semenava I.; manuscript editing and approval of the final version for publication – Marchenko L. and Fedulov A.

Submitted: 23.03.2026

Accepted: 11.05.2026

Contacts: tvk35@Yahoo.com

Abstract

Purpose. To comprehensively assess clinical and structural predictors of visual acuity in patients with age-related macular degeneration (AMD), develop a predictive model for the risk of low vision, and identify disease subtypes using machine learning methods.

Materials and methods. A retrospective analysis of data from 463 eyes (278 patients) with AMD was performed. Demographic, functional (best-corrected visual acuity, BCVA), and structural optical coherence tomography parameters were evaluated. The Kruskal – Wallis test was used for group comparisons. The Random Forest algorithm was applied to rank predictor importance, logistic regression to predict the risk of low vision ($BCVA < 0.1$), and K-means clustering to identify phenotypes.

Results. The most significant predictors of BCVA were AREDS stage (0.164), RPE status (0.110), and central foveolar retinal thickness (CFRT; 0.106). The logistic regression model showed good predictive performance for low vision ($AUC=0.81$); key risk factors included increased average retinal volume ($OR=2.24$) and the presence of fibrovascular detachment ($OR=1.98$). Cluster analysis revealed three phenotypes: active neovascularization (moderate vision loss, high frequency of exudative changes), edematous/fibrotic (most severe, with maximal retinal thickness and worst functional outcomes), and drusen/early AMD (relatively preserved vision, high drusen frequency).

Conclusion. The application of machine learning methods confirmed the critical role of neurosensory retinal structural integrity and identified three clinically relevant AMD phenotypes. The developed predictive model and phenotyping approach can inform personalized monitoring and treatment strategies for AMD patients.

Keywords: age-related macular degeneration, machine learning, random forest, clustering, visual acuity, prognosis, phenotypes, optical coherence tomography

■ ВВЕДЕНИЕ

Возрастная макулярная дегенерация (ВМД) остается ведущей причиной необратимого снижения центрального зрения у лиц старше 50 лет в развитых странах, оказывая существенное влияние на качество жизни пациентов и создавая значительную медико-социальную нагрузку [1, 2]. Заболевание характеризуется гетерогенностью клинических проявлений и темпов прогрессирования: от медленно развивающихся друз и изменений пигментного эпителия до быстропрогрессирующей неоваскуляризации и фиброза, приводящих к необратимой потере зрительных функций [3, 4]. Ключевым функциональным показателем, определяющим степень инвалидизации пациента, является максимальная скорректированная острота зрения (МКОЗ). Понимание факторов, ассоциированных с ее снижением, и выделение клинически однородных подгрупп пациентов имеет первостепенное значение для разработки персонализированных стратегий наблюдения и лечения [5]. Традиционные статистические подходы, такие как сравнение средних и корреляционный анализ, позволяют оценить вклад отдельных параметров, однако они ограничены в способности учитывать сложные нелинейные взаимодействия между множеством клинических, демографических и структурных признаков [6]. Между тем современные офтальмологические исследования генерируют обширные массивы данных, включающие как количественные показатели (возраст, толщина сетчатки, объемные характеристики), так и качественные (наличие отслоек, состояние пигментного эпителия и фоторецепторов). Интеграция этих разнородных данных для построения прогностических моделей и выявления скрытых закономерностей требует применения более сложных аналитических инструментов [7]. В последние годы методы машинного обучения (МО) все активнее внедряются в офтальмологию, демонстрируя высокую эффективность в анализе медицинских изображений и структурированных клинических данных [8, 9]. В отличие от классической статистики, алгоритмы МО способны одновременно обрабатывать большое число предикторов, выявлять неочевидные взаимосвязи и строить модели с высокой предсказательной способностью. Случайный лес (Random Forest) позволяет ранжировать признаки по их вкладу в целевую переменную, что дает возможность выделить ключевые факторы, определяющие остроту зрения [10]. Логистическая регрессия остается «золотым стандартом» для бинарной классификации (например, прогноза низкого зрения) благодаря своей интерпретируемости, позволяя оценить отношение шансов для каждого предиктора [11]. Наконец, методы кластеризации (в частности K-средних) относятся к обучению без учителя и дают возможность обнаружить естественные группировки пациентов на основе всего комплекса признаков, что может способствовать идентификации новых фенотипов ВМД, не определяемых клинически [12, 13].

■ ЦЕЛЬ ИССЛЕДОВАНИЯ

Провести комплексную оценку клинических и структурных предикторов остроты зрения у пациентов с ВМД, построить прогностическую модель риска низкой остроты зрения и выделить клинически значимые подтипы заболевания с использованием методов многомерного и интегративного анализа данных.



■ МАТЕРИАЛЫ И МЕТОДЫ

В ретроспективное исследование были включены данные 463 глаз 278 пациентов с подтвержденным диагнозом возрастной макулярной дегенерации. Критерии включения: наличие данных первичного обследования до начала какого-либо этиотропного лечения, полный набор клинических и томографических параметров, необходимых для анализа. Из исследования исключались глаза с сопутствующей офтальмологической патологией, влияющей на остроту зрения (выраженная катаракта, глаукома, диабетическая ретинопатия), а также случаи с отсутствием ключевых переменных после предварительной обработки данных.

Клинические и инструментальные параметры

Регистрировались следующие показатели: демографические – возраст (полных лет), пол; функциональные – максимальная скорректированная острота зрения (МКОЗ); структурные – центральная толщина сетчатки в фовеоле (ЦТС foveola), толщина сетчатки в парафовеолярной области (ЦТС sup. inner fovea и ЦТС sup. out. fovea), общий объем сетчатки (total volume) и средний объем (average volume) по данным оптической когерентной томографии (ОКТ); статусные (категориальные) – стадия по AREDS (Age-Related Eye Disease Study, 1–4), состояние ретинального пигментного эпителия (RPE status: 1 – сохранен, 2 – неравномерный, 3 – единичные разрывы, 4 – множественные разрывы, 5 – не определяется), состояние эллипсоидной зоны (EZ status: 1 – сохранена, 2 – фрагментация, 3 – локально не определяется, 4 – не определяется), состояние миоидной зоны (MyoidnZ status: аналогичные градации); бинарные признаки, отражающие наличие патологических изменений – кистозный макулярный отек (CME localisation_bin), серозная отслойка РПЭ (Serous RPE detachment localisation_bin), геморрагическая отслойка РПЭ (Hemorrhagic RPE detachment localisation_bin), фиброваскулярная отслойка РПЭ (Fibrovascular RPE detachment localisation_bin), друзеноидная отслойка РПЭ (Drusenoid detachment RPE localisation_bin), друзы (Druses localisation_bin), жидкость под РПЭ (Fluid under RPE localisation_bin), отслойка нейроэпителия (RNE detachment localisation_bin), гиперрефлективный материал (гиперрефлективный материал локализация_bin). При построении моделей все бинарные переменные кодировались как 1 при наличии признака (значение >0 в исходной переменной) и 0 при его отсутствии.

Предобработка данных

Пропуски в данных были минимальными (<0.5%) и заполнялись медианным значением для непрерывных переменных. Для устранения влияния масштаба на результаты регрессионного анализа и кластеризации все количественные признаки были стандартизированы (центрирование и нормирование к единичному стандартному отклонению).

Статистический анализ

На первом этапе проводилось сравнение демографических и клинических характеристик между тремя группами, выделенными на основе МКОЗ: высокая острота зрения ($\text{МКОЗ} \geq 0,5$), средняя ($0,1 \leq \text{МКОЗ} < 0,5$) и низкая ($\text{МКОЗ} < 0,1$). Для количественных переменных применяли критерий Краскела – Уоллеса с последующими попарными сравнениями с помощью U-критерия Манна – Уитни и поправкой Бонферрони.

Для категориальных переменных использовали критерий χ^2 (или точный критерий Фишера при малых ожидаемых частотах). Статистически значимыми считали различия при $p < 0,05$.

Методы машинного обучения

Для решения трех взаимосвязанных задач – выявление наиболее информативных признаков, прогнозирование низкой остроты зрения и выделение скрытых подтипов ВМД – были применены различные алгоритмы машинного обучения (МО).

Оценка важности признаков

Проводилась с помощью алгоритма случайного леса (Random Forest) [10]. Определяли вклад каждого предиктора в предсказание МКОЗ. Модель состояла из 100 деревьев; для оценки важности использовалось усредненное уменьшение примесей (Gini impurity). Полученный рейтинг позволил отобрать переменные для последующего моделирования.

Прогнозирование низкого зрения

Для бинарной классификации (исход: МКОЗ $< 0,1$) использовали логистическую регрессию с L2-регуляризацией. Данные разделялись на обучающую (70%) и тестовую (30%) выборки с сохранением пропорции классов. Модель обучалась на стандартизированных признаках. Качество классификации оценивали по площади под ROC-кривой (AUC-ROC), чувствительности и специфичности. Для интерпретации влияния предикторов рассчитывали отношения шансов (OR) с 95% доверительными интервалами. Выбор порога классификации (0,5) был обусловлен задачей сохранения интерпретируемости модели и стандартным подходом к бинарной классификации. Методы балансировки классов (например, Synthetic Minority Over-sampling Technique (SMOTE)) не применялись, чтобы избежать искажения коэффициентов и риска переобучения на несбалансированной выборке.

Кластеризация

Для выявления естественных группировок пациентов на основе всего набора клинических и структурных признаков применялся метод К-средних (K-means) [12]. Оптимальное число кластеров выбирали, исходя из клинической интерпретируемости, поскольку традиционные критерии (метод локтя, силуэтный коэффициент) не давали однозначного решения. Значения индекса силуэта (silhouette score) для $k=2, 3, 4$ составили 0,32, 0,29 и 0,24 соответственно, что не позволило сделать однозначный выбор, поэтому предпочтение было отдано клинической интерпретируемости. Кластеризация проводилась на стандартизированных данных с использованием евклидовой метрики (кластеризация – это метод, который автоматически группирует пациентов со схожими характеристиками, без заранее заданных критериев). Для визуализации полученных кластеров применяли метод главных компонент, Principal Component Analysis (PCA) с проекцией на первые две компоненты. На полученном графике каждый пациент представлен точкой; близко расположенные точки соответствуют пациентам со сходными характеристиками, что позволяет наглядно оценить степень разделения выделенных фенотипов. Все расчеты выполнены в среде с использованием библиотек pandas, numpy, scikit-learn, scipy, matplotlib и seaborn.



■ РЕЗУЛЬТАТЫ

Общая клинико-демографическая характеристика пациентов представлена в табл. 1. Средний возраст включенных в исследование пациентов с ВМД составил $73,5 \pm 8,2$ года.

Представленная выборка отражает типичный контингент пациентов с ВМД: преобладают женщины старше 70 лет, у большинства диагностирована поздняя (4-я) стадия заболевания. Высокая частота структурных изменений (фиброваскулярные отслойки, гиперрефлективный материал) указывает на активный неоваскулярный процесс. Эти данные служат основой для дальнейшего анализа факторов, определяющих остроту зрения. Распределение пациентов по уровню остроты зрения: на основе исходной МКОЗ пациенты были разделены на три группы – высокая острота зрения ($\text{МКОЗ} \geq 0,5$; $n=202$, 43,6%), средняя острота зрения ($0,1 \leq \text{МКОЗ} < 0,5$; $n=188$, 40,6%) и низкая острота зрения ($\text{МКОЗ} < 0,1$; $n=73$, 15,8%). Сравнительная характеристика групп представлена в табл. 2.

Из полученных данных следует, что ухудшение остроты зрения закономерно коррелирует с увеличением возраста, толщины сетчатки и частоты деструктивных изменений (повреждение РПЭ, отсутствие эллипсоидной зоны, фиброваскулярные отслойки). Интересно, что наличие друз, напротив, ассоциировано с более высоким зрением, что соответствует их роли как маркера ранних стадий ВМД. Эти закономерности, выявленные с помощью классического статистического анализа, задают направления для более глубокого машинного обучения.

Оценка важности признаков (Random Forest): для определения наиболее значимых предикторов остроты зрения применен алгоритм случайного леса. Ранжированный список признаков по их вкладу в предсказание МКОЗ представлен на рис. 1. Наибольшую важность продемонстрировали: стадия по AREDS (0,164) – интегральный показатель тяжести ВМД, отражающий совокупность структурных изменений, состояние пигментного эпителия (RPE status, 0,110) – повреждение РПЭ критически нарушает трофику фоторецепторов; центральная толщина сетчатки (ЦТС foveola, 0,106) – увеличение толщины может свидетельствовать об отеке, фиброзе или отслойке; толщина в верхнем отделе (ЦТС sup. out. fovea; 0,099) – отражает распространенность изменений за пределы фовеа; возраст (0,094) – естественный фактор

Таблица 1
Клинико-демографическая характеристика пациентов с ВМД (n=463)
Table 1
Clinical and demographic profile of AMD patients (n=463) (n=463)

Параметр	Значение
Возраст, лет ($M \pm SD$)	$73,5 \pm 8,2$
Пол (женщины/мужчины), n (%)	304 (65,7%) / 159 (34,3%)
МКОЗ ($M \pm SD$; медиана)	$0,41 \pm 0,29$; 0,40
Длительность заболевания, лет ($M \pm SD$)	$2,3 \pm 1,9$
ЦТС foveola, мкм ($M \pm SD$)	$274,1 \pm 77,8$
Стадия AREDS 4, n (%)	320 (69,1%)
Наличие кистозного отека (CME), n (%)	80 (17,3%)
Наличие фиброваскулярной отслойки, n (%)	157 (33,9%)
Наличие друз, n (%)	237 (51,2%)
Наличие гиперрефлективного материала, n (%)	206 (44,5%)

Таблица 2

Сравнительная характеристика групп с разным уровнем МКОЗ ($M \pm SD$ – среднее $\pm$ стандартное отклонение; p – критерий Краскела – Уоллиса для количественных и χ^2 для категориальных переменных)

Table 2

Comparative characteristics of groups with different BCVA levels ($M \pm SD$ – mean $\pm$ standard deviation; p – Kruskal – Wallis test for quantitative variables and χ^2 test for categorical variables)

Параметр	Высокая МКОЗ (n=202)	Средняя МКОЗ (n=188)	Низкая МКОЗ (n=73)	p
Возраст, лет	71,6 $\pm$ 7,9	74,7 $\pm$ 8,1	75,6 $\pm$ 8,5	<0,001
ЦТС foveola, мкм	248,4 $\pm$ 62,1	285,6 $\pm$ 75,3	311,7 $\pm$ 101,2	<0,001
Average volume, мкм ³	269,2 $\pm$ 20,1	275,6 $\pm$ 26,3	291,5 $\pm$ 51,2	0,003
Стадия AREDS 4, n (%)	96 (47,5%)	152 (80,9%)	72 (98,6%)	<0,001
RPE status 5 (не определяется), n (%)	13 (6,4%)	39 (20,7%)	54 (74,0%)	<0,001
EZ status 4 (не определяется), n (%)	66 (32,7%)	108 (57,4%)	66 (90,4%)	<0,001
Кистозный отек, n (%)	11 (5,4%)	43 (22,9%)	26 (35,6%)	<0,001
Фиброваскулярная отслойка, n (%)	27 (13,4%)	76 (40,4%)	48 (65,8%)	<0,001
Друзы, n (%)	136 (67,3%)	87 (46,3%)	11 (15,1%)	<0,001

прогрессирования ВМД, толщина во внутреннем отделе (ЦТС sup. inner fovea; 0,083); состояние эллипсоидной зоны (EZ status; 0,069) – маркер целостности фоторецепторов, ее повреждение напрямую снижает остроту зрения; средний объем сетчатки (average volume; 0,061) – коррелирует с общим объемом патологических изменений; состояние миоидной зоны (MyoidnZ status; 0,058) – отражает нарушения во внутренних сегментах фоторецепторов; общий объем сетчатки (total volume; 0,046) – менее информативен, чем средний объем; длительность заболевания (Длитзабол_лет; 0,022) – влияет, но слабее структурных параметров. Бинарные признаки наличия патологических изменений (гиперрефлективный материал, отслойки, кистозный отек, друзы) имели существенно меньший вес (от 0,016 до 0,0003), что указывает на то, что их влияние уже опосредовано через количественные и порядковые показатели (толщину, объем, статусы).

Случайный лес объективно подтвердил ключевую роль структурной целостности нейросенсорной сетчатки в сохранении зрительных функций. Особого внимания заслуживает высокая важность состояния эллипсоидной и миоидной зон – маркеров жизнеспособности фоторецепторов. Их повреждение, часто необратимое, оказывается более значимым для прогноза зрения, чем многие другие факторы. Несколько меньший вес EZ status по сравнению с RPE status может объясняться тесной патофизиологической связью между этими структурами: повреждение РПЭ, как правило, влечет за собой нарушение функции и целостности фоторецепторов, и модель могла «отдать предпочтение» более интегральному «показателю». Относительно низкая важность бинарных признаков осложнений объясняется тем, что их влияние уже «поглощено» количественными параметрами (например, фиброваскулярная отслойка увеличивает толщину и объем). Таким образом, для практического прогнозирования целесообразно ориентироваться на интегральные показатели: стадию AREDS, состояние РПЭ и эллипсоидной зоны, а также центральную толщину сетчатки.

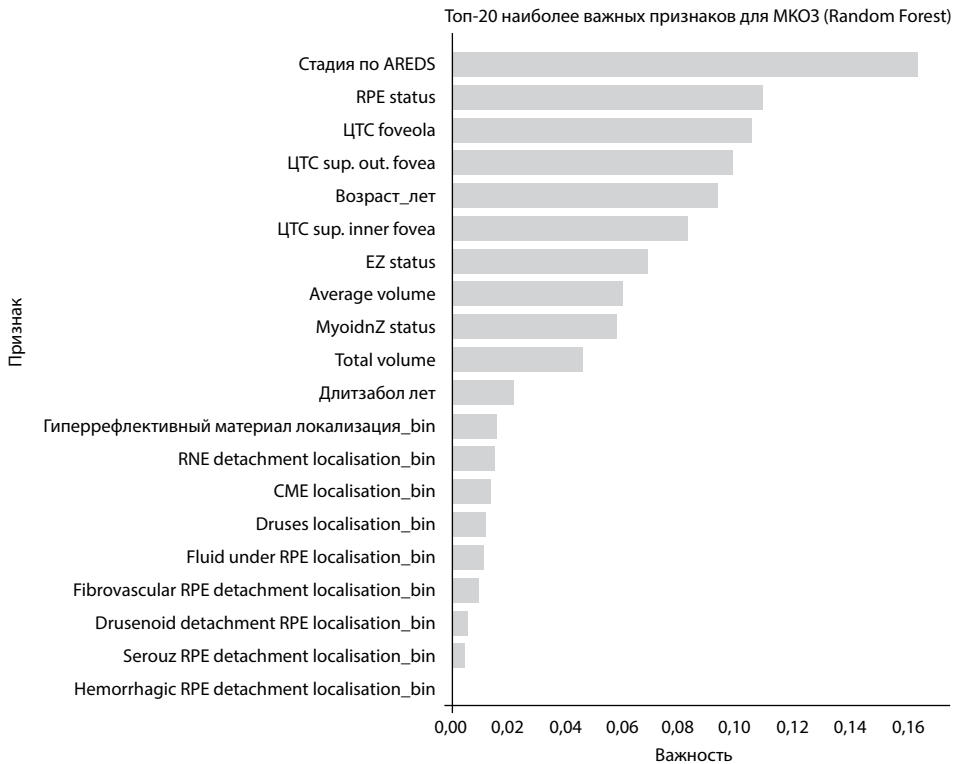


Рис. 1. Важность признаков для предсказания МКОЗ (метод случайного леса)
Fig. 1. Feature importance for predicting BCVA (Random Forest method)

Прогнозирование низкого зрения (логистическая регрессия)

Модель бинарной логистической регрессии для предсказания низкого зрения (МКОЗ $<0,1$) продемонстрировала хорошую дискриминационную способность: площадь под ROC-кривой на тестовой выборке составила 0,813 (95% ДИ 0,72–0,90). ROC-кривая представлена на рис. 2. При пороге классификации 0,5 чувствительность модели составила 27%, специфичность – 98%.

Модель с высокой специфичностью (98%) позволяет надежно идентифицировать пациентов, у которых низкое зрение маловероятно, $AUC=0,81$ свидетельствует о хорошем качестве модели, которая может быть использована для первичного скрининга групп риска.

Коэффициенты модели и соответствующие отношения шансов (OR) приведены в табл. 3. Увеличение среднего объема сетчатки (average volume), наличие фиброваскулярной отслойки, возраст, увеличение ЦТC foveola, кистозный отек и наличие гиперрефлективного материала ассоциировались с повышенным риском низкого зрения. Напротив, увеличение толщины в парафовеолярной области (ЦТC sup. inner fovea, ЦТC sup. out. fovea) и наличие друз снижали этот риск.

Логистическая регрессия, будучи интерпретируемым методом, позволила количественно оценить вклад каждого фактора. Наибольшее повышение шансов низкого зрения связано с увеличением среднего объема сетчатки ($OR=2,24$), что может

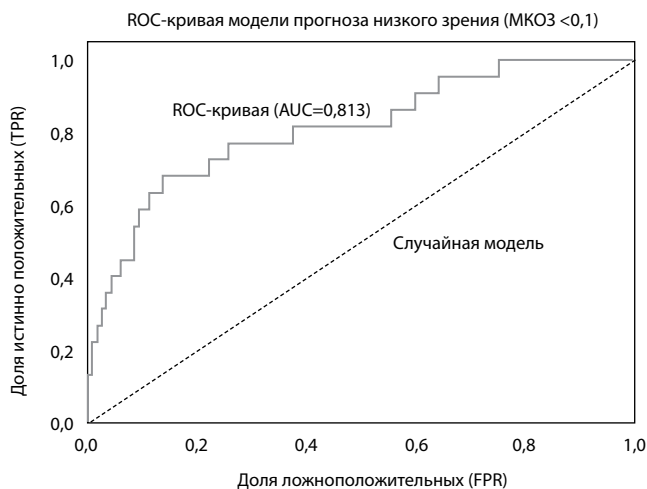


Рис. 2. ROC-кривая модели прогноза низкого зрения, AUC=0,813; пунктирная линия – случайная модель
Fig. 2. ROC curve of the low vision prediction model, AUC=0.813; the dotted line represents a random model

отражать наличие отека или фиброза. Защитный эффект друз (OR=0,58) и увеличения толщины в парафовеолярной зоне подтверждает их связь с более ранними и менее агрессивными формами ВМД. Эти результаты могут быть использованы для разработки клинического калькулятора риска.

Кластеризация пациентов (К-средние)

Методом К-средних при k=3 выделены три кластера, клиническая характеристика которых представлена в табл. 4.

На рис. 3 показано распределение пациентов в пространстве первых двух главных компонент.

Таблица 3
Коэффициенты и отношения шансов в модели логистической регрессии
Table 3
Coefficients and odds ratios in the logistic regression model

Признак	Коэффициент (β)	Отношение шансов (OR)	95% ДИ	p
Average volume	0,807	2,24	1,48–3,41	<0,001
Фиброваскулярная отслойка	0,488	1,63	1,12–2,38	0,010
Возраст	0,383	1,47	1,08–1,99	0,014
ЦТС foveola	0,349	1,42	1,03–1,96	0,032
Кистозный отек	0,260	1,30	0,94–1,79	0,112
Гиперрефлективный материал	0,250	1,28	0,91–1,81	0,156
ЦТС sup. inner fovea	–0,381	0,68	0,46–1,01	0,054
ЦТС sup. out. fovea	–0,544	0,58	0,39–0,87	0,008
Друзы	–0,548	0,58	0,39–0,86	0,007



Таблица 4
Характеристика кластеров пациентов с ВМД
Table 4
Characteristics of AMD patient clusters

Параметр	Кластер 0 (n=153)	Кластер 1 (n=61)	Кластер 2 (n=249)
Возраст, лет (M±SD)	73,3±8,2	75,3±8,5	73,2±8,1
МКОЗ (медиана)	0,3 (0,2)	0,2 (0,15)	0,6 (0,6)
ЦТС foveola, мкм	271±72	410±84	241±51
Average volume, мкм ³	272±25	327±48	265±18
Кистозный отек, %	24	59	4
Фиброваскулярная от- слойка, %	71	57	3
Друзы, %	10	18	84
Гиперрефлективный материал, %	80	59	19
Стадия AREDS 4, %	98,7	98,4	43,8
Низкое зрение (<0,1), %	16,3	41,0	8,4

Каждая точка соответствует одному глазу. Различные типы маркеров обозначают принадлежность к одному из трех кластеров. Проекция на первые две главные компоненты демонстрирует относительную обособленность выделенных групп.

Кластеризация, выполненная без априорных гипотез, выделила три клинических фенотипа ВМД. Кластер 0 (активная неоваскуляризация) – пациенты с 4-й стадией, высокой активностью процесса (фиброваскулярные отслойки, гиперрефлективный материал) и умеренным снижением зрения. Кластер 1 (отечный/фиброзный) – наиболее тяжелый вариант с максимальной толщиной сетчатки, частым отеком и самым низким зрением; именно здесь сконцентрирован высокий риск неблагоприятного исхода (41% пациентов с МКОЗ <0,1). Кластер 2 (друзный/ранний) – характеризуется



Рис. 3. Визуализация кластеров методом главных компонент (PCA)
Fig. 3. Visualization of clusters using principal component analysis (PCA)

сохранным зрением, низкой частотой осложнений и высокой долей друз; прогностически благоприятный. Такая стратификация может служить основой для дифференцированного подхода к лечению и наблюдению.

■ ОБСУЖДЕНИЕ

В настоящем исследовании с использованием методов машинного обучения проведен комплексный анализ клинических и структурных факторов, ассоциированных с остротой зрения у пациентов с ВМД. Основными результатами стали: 1) ранжирование предикторов МКОЗ с выделением наиболее значимых – стадии AREDS (Age-Related Eye Disease Study), состояния пигментного эпителия и центральной толщины сетчатки; 2) построение прогностической модели риска низкого зрения (МКОЗ $<0,1$) с хорошей дискриминационной способностью ($AUC=0,81$); 3) идентификация трех клинических фенотипов ВМД методом кластеризации. Полученные данные не только подтверждают известные закономерности, но и предлагают новые подходы к стратификации пациентов. Высокая прогностическая значимость стадии AREDS ожидаема и согласуется с многочисленными исследованиями, подтверждающими, что прогрессирование ВМД от ранних к поздним стадиям сопровождается необратимым снижением зрения [3, 4]. Однако применение случайного леса позволило количественно оценить вклад других факторов в условиях мультиколлинеарности. Обращает на себя внимание, что состояние пигментного эпителия (RPE status) оказалось более важным предиктором, чем состояние эллипсоидной зоны (EZ status). Это не противоречит клиническим представлениям, а скорее отражает патофизиологическую иерархию: первичное повреждение РПЭ нарушает трофику фоторецепторов, что вторично ведет к дегенерации эллипсоидной зоны. В модели, включающей оба коррелирующих признака, алгоритм мог «выбрать» RPE status как более интегральный маркер, аккумулирующий информацию о состоянии фоторецепторов. Тем не менее значимость EZ status (0,069) остается высокой и подчеркивает необходимость оценки целостности фоторецепторов при прогнозировании исходов. Сходные результаты были получены в работах, где комбинация структурных и функциональных параметров улучшала предсказание прогрессирования ВМД [7, 9]. Центральная толщина сетчатки (ЦТС foveola) и средний макулярный объем вошли в число ведущих количественных предикторов. Увеличение этих показателей, вероятно, отражает наличие интравитреальной жидкости, фиброваскулярной ткани или отека, что напрямую ухудшает зрение. Интересно, что наличие друз, напротив, ассоциировалось с более высокой остротой зрения и меньшим риском низкой остроты зрения, что соответствует их роли как маркера ранних стадий [5]. Аналогичные закономерности были описаны в исследованиях, использующих кластерный анализ для выделения фенотипов ВМД [12, 13].

Прогностическая модель низкого зрения

Полученная модель логистической регрессии продемонстрировала высокую специфичность. Для повышения чувствительности в будущих исследованиях могут быть применены методы балансировки классов, такие как SMOTE, или оптимизация порога классификации. В рамках настоящей работы мы сознательно не использовали эти подходы, чтобы сохранить интерпретируемость коэффициентов модели и избежать риска переобучения на несбалансированной выборке. Кроме того, выбор



порога классификации требует клинической валидации и учета соотношения пользы и вреда от ложноположительных и ложноотрицательных решений, что выходит за рамки данного исследования. Тем не менее $AUC=0,81$ свидетельствует о хорошем качестве модели, сопоставимом с опубликованными прогностическими инструментами для ВМД [9, 11]. Полученные отношения шансов (табл. 3) позволяют количественно оценить вклад каждого фактора: например, увеличение среднего объема на одно стандартное отклонение повышает шансы низкого зрения в 2,24 раза. Такие данные могут лечь в основу клинического калькулятора для выделения групп риска, требующих более частого мониторинга или раннего лечения.

Кластеризация и выделение фенотипов ВМД

Применение метода К-средних без априорных гипотез позволило выделить три клинически интерпретируемых кластера. Кластер 0 (активная неоваскуляризация) – пациенты с поздней стадией, высокой частотой фиброваскулярных отслоек и гиперрефлективного материала, но с умеренным снижением зрения. Вероятно, это группа с активной неоваскуляризацией, где возможно улучшение на фоне анти-VEGF-терапии. Кластер 1 (отечный/фиброзный) – наиболее тяжелый вариант, характеризующийся максимальной толщиной сетчатки, частым кистозным отеком и самым низким зрением. Пациенты этого кластера имеют наибольший риск неблагоприятного исхода (41% с МКОЗ $<0,1$) и, возможно, нуждаются в более агрессивной тактике. Кластер 2 (друзный/ранний) – включает пациентов с разнообразными стадиями, высокой частотой друз и относительно сохранным зрением, что соответствует благоприятному прогнозу. Эти фенотипы согласуются с известными клиническими формами ВМД [4, 13] и могут служить основой для персонализированного подхода. Полученные нами фенотипы также перекликаются с результатами недавнего исследования Holland с соавт. [14], где с помощью методов глубокого обучения на основе ОКТ-изображений были выделены кластеры биомаркеров, ассоциированные с различной скоростью прогрессирования ВМД. Это подтверждает перспективность использования алгоритмов кластеризации для выявления гетерогенности заболевания и подчеркивает важность интеграции структурных биомаркеров в прогностические модели. Наше исследование демонстрирует, что алгоритмы машинного обучения способны не только воспроизводить известные клинические закономерности, но и выявлять скрытые взаимосвязи, ранжировать факторы по значимости и выделять новые фенотипы. В отличие от классической статистики эти методы учитывают нелинейные взаимодействия и могут обрабатывать большое число переменных без априорных предположений. Полученные результаты могут быть использованы для разработки цифровых инструментов поддержки принятия решений, где врач вводит параметры пациента и получает оценку риска низкого зрения или вероятную принадлежность к фенотипу. Это особенно актуально в условиях персонализированной медицины, где выбор тактики лечения должен основываться на комплексной оценке всех доступных данных. Таким образом, применение методов машинного обучения позволило углубить понимание факторов, определяющих остроту зрения при ВМД, построить интерпретируемую прогностическую модель и выделить три клинических фенотипа. Полученные данные развивают результаты нашего предыдущего исследования, посвященного использованию алгоритмов машинного обучения в офтальмологии как инструмента поддержки принятия клинических решений [15]. В совокупности эти

работы демонстрируют потенциал интеграции многомерного анализа данных в рутинную клиническую практику для персонализации ведения пациентов с ВМД.

■ ВЫВОДЫ

1. Методы машинного обучения (случайный лес, логистическая регрессия, кластеризация) позволили выявить ключевые предикторы остроты зрения при ВМД: стадия AREDS, состояние пигментного эпителия и центральная толщина сетчатки.
2. Разработанная модель прогноза низкого зрения (МКОЗ <0,1) показала хорошее качество (AUC=0,81); основными факторами риска являются увеличение объема сетчатки, фиброваскулярная отслойка и возраст.
3. Кластерный анализ выделил три клинических фенотипа ВМД: активная неоваскуляризация, отечный/фиброзный (наиболее тяжелый) и друзный/ранний (благоприятный), что открывает возможности для персонализированного подхода.
4. Продемонстрирована эффективность интеграции методов ИИ в анализ офтальмологических данных для стратификации пациентов и прогнозирования исходов.

■ ЛИТЕРАТУРА/REFERENCES

1. Plyaskina US, Fomina AV. Epidemiological aspects of age-related macular degeneration and senile cataract in Russia and in the world. *Sovremennye problemy zdravookhraneniya i meditsinskoj statistiki*. 2024;3:713–730. DOI: 10.24412/2312-2935-2024-3-713-730
2. Klein R, Klein BE, Knudtson MD, et al. Prevalence of age-related macular degeneration in the US population. *Arch Ophthalmol*. 2011;129(1):75–80. DOI: 10.1001/archophthol.2010.318
3. Age-Related Eye Disease Study Research Group. The Age-Related Eye Disease Study (AREDS): design implications. AREDS report no. 1. *Control Clin Trials*. 1999;20(6):573–600. DOI: 10.1016/s0197-2456(99)00031-8
4. Ferris FL 3rd, Wilkinson CP, Bird A, et al. Clinical classification of age-related macular degeneration. *Ophthalmology*. 2013;120(4):844–851. DOI: 10.1016/j.ophtha.2012.10.036
5. Age-Related Eye Disease Study Research Group. A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E, beta carotene, and zinc for age-related macular degeneration and vision loss: AREDS report no. 8. *Arch Ophthalmol*. 2001;119(10):1417–1436. DOI: 10.1001/archophth.119.10.1417
6. Chistyakov SP. Random forests: a review. *Trudy Karel'skogo nauchnogo tsentra RAN*. 2013;1:117–136. (In Russ.)
7. Schmidt-Erfurth U, Bogunović H, Sadeghipour A, et al. Machine learning to analyze the prognostic value of imaging biomarkers in neovascular AMD. *Sci Rep*. 2020;10(1):12345. DOI: 10.1038/s41598-020-69121-7
8. Ting DS, Peng L, Varadarajan AV, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol*. 2019;103(2):167–175. DOI: 10.1136/bjophthalmol-2018-313173
9. Kuryшева NI, Pomerantsev AL, Rodionova OY, et al. [Results of using machine learning to determine predictors of hypotensive efficacy of peripheral laser iridotomy in patients with primary angle closure]. *Oftalmokhirurgiya*. 2025;2:XX–XX. (In Russ.). DOI: 10.25276/0235-4160-2025-2-XX
10. Kuryшева NI, Sharova GA. Suspicion of primary angle closure: application of machine learning to substantiate close monitoring. *Vestnik oftalmologii*. 2025;141(2):67–74. (In Russ.). DOI: 10.17116/oftalma202514102167
11. Kuryшева NI, Pomerantsev AL, Rodionova OY, et al. Comparison of lens extraction versus laser iridotomy on anterior segment, choroid and intraocular pressure in primary angle closure using machine learning. *J Glaucoma*. 2023;32(6):e43–e55. DOI: 10.1097/IJG.0000000000002145
12. Zagidullina AS, Aznabaev BM, Zulkarnaev RH, et al. Application of cluster analysis in the formation of primary open-angle glaucoma phenotypes. *Meditsinskiy vestnik Bashkortostana*. 2022;17(4):5–9. (In Russ.)
13. Waldstein SM, Vogl WD, Bogunovic H, et al. Unsupervised clustering identifies clinically relevant phenotypes in age-related macular degeneration. *Retina*. 2021;41(3):489–497. DOI: 10.1097/IAE.0000000000002894
14. Holland R, Kaye R, Hagag AM, et al. Deep Learning – Based Clustering of OCT Images for Biomarker Discovery in Age-Related Macular Degeneration (PINNACLE Study Report 4). *Ophthalmology Science*. 2024;100543. DOI: 10.1016/j.xops.2024.100543
15. Kachan TV, Skrypnik OV, Kurochkin AV, et al. Opyt ispol'zovaniya metodov mashinnogo obucheniya kak instrumenta podderzhki prinjatija klinicheskikh reshenij v oftalmologii. *Medicinskie novosti*. 2023;6:27–31. (In Russ.)