

УДК 616-092

Новые подходы к диагностике неходжкинских В-клеточных лимфом и хронического лимфолейкоза по кластерам дифференцировки нормальных и опухолевых лимфоцитов периферической крови

Черкашенко В.Н.¹, Сабурин И.Н.¹, Кузьмина Е.Г.², Антонов К.Н.³, Устинова Е.Е.¹, Щипкова Е.С.¹, Зацаренко С.В.², Мушкарина Т.Ю.², Меркулова Д.М.⁴, Морозов С.Г.¹

¹ Федеральное государственное бюджетное научное учреждение «Научно-исследовательский институт общей патологии и патофизиологии» 125315, Москва, ул. Балтийская, д. 8

² Медицинский радиологический научный центр имени А.Ф. Цыба – филиал Федерального государственного бюджетного учреждения «Научный медицинский исследовательский центр радиологии» Министерства здравоохранения Российской Федерации 249036, Калужская область, Обнинск, ул. Королёва, д. 4

³ Факультет вычислительной математики и кибернетики Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет имени М.В. Ломоносова» 119234, Москва, Ленинские горы, д. 1, стр. 52

⁴ Частное учреждение здравоохранения Центральная клиническая больница «РЖД-Медицина» 129128, Москва, Будайская ул., д. 2

Актуальность. В настоящее время быстро развиваются и внедряются технологии машинного обучения и искусственного интеллекта в научные исследования, что значительно упрощает изучение и использование прикладных разработок в медицине. Разработка метода автоматической классификации неходжкинских В-клеточных лимфом (НХЛ) и хронического лимфолейкоза (ХЛЛ) в сравнении с нормой (Н) позволит проанализировать различия многомерных взаимосвязей в системе «иммунитет – опухолевый рост» и создаст возможность классифицировать эти заболевания по данным крови, не используя для диагностики материалы других органов и тканей.

Цель исследования – создать автоматизированную классификацию НХЛ и ХЛЛ в сравнении с нормой и оценить качество диагностики по различиям показателей многомерных образов иммунокомпетентных и опухолевых клеток в периферической крови.

Материалы и методы. Проведено сравнительное исследование иммунитета и опухолевых клеток в периферической крови 185 здоровых людей, 352 пациентов с В-клеточной НХЛ и 315 пациентов с ХЛЛ. Методом машинной обработки, классифицирующим данные путем построения дерева решений (ДР), сопоставлены показатели многомерных образов, формируемых субпопуляциями лимфоцитов и опухолевыми клетками в периферической крови. Построены 3 классификатора состояний: НХЛ в сравнении с нормой (Н–НХЛ), ХЛЛ в сравнении с нормой (Н–ХЛЛ) и НХЛ в сравнении с ХЛЛ (НХЛ–ХЛЛ). Качество классификаторов определено по критериям матрицы ошибок и ROC-кривой.

Результаты. Построенные методом ДР классификаторы значений и характера распределения показателей иммунитета и опухолевых клеток различают состояния Н–НХЛ, Н–ХЛЛ и НХЛ–ХЛЛ и определяют алгоритм классификации. Критериями методов матрицы ошибок и ROC-кривой подтверждено высокое качество классификаторов.

Заключение. Методом машинного обучения построено ДР, классифицирующее НХЛ и ХЛЛ, что позволяет реализовать альтернативную диагностику НХЛ и ХЛЛ, используя только данные о наличии иммунных и опухолевых лимфоцитов в периферической крови.

Ключевые слова: периферическая кровь; иммунитет; опухолевые лимфоциты; неходжкинские В-клеточные лимфомы; хронический лимфолейкоз; машинная классификация; алгоритм диагностики.

Для цитирования: Черкашенко В.Н., Сабурин И.Н., Кузьмина Е.Г., Антонов К.Н., Устинова Е.Е., Щипкова Е.С., Зацаренко С.В., Мушкарина Т.Ю., Меркулова Д.М., Морозов С.Г. Новые подходы к диагностике неходжкинских В-клеточных лимфом и хронического лимфолейкоза по кластерам дифференцировки нормальных и опухолевых лимфоцитов периферической крови. *Патогенез.* 2024; 22(4): 64–77.

DOI: 10.25557/2310-0435.2024.04.64-77

Для корреспонденции: Черкашенко Владимир Николаевич, e-mail: comrade_cherkashenko@mail.ru

Финансирование. Исследование не имеет спонсорской поддержки

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила: 05.12.2024.

New approaches to the diagnosis of Non-Hodgkin's B-cell lymphomas and Chronic Lymphocytic Leukemia by clusters of differentiation of normal and tumor lymphocytes in peripheral blood

Cherkashenco V.N.¹, Saburina I.N.¹, Kuzmina E.G.², Antonov K.N.³, Ustinova E.E.¹, Shchipkova E.S.¹, Zacarenko S.V.², Mushkarina T.Yu.², Merkulova D.M.⁴, Morozov S.G.¹

¹ Institute of General Pathology and Pathophysiology
Baltiyskaya Str. 8, Moscow 125315, Russian Federation

² A. Tsyb Medical Radiological Research Center – branch of the National Medical Research Radiological Center
Korolev Str. 4, Obninsk of Kaluga Region 249036, Russian Federation

³ M.V. Lomonosov Moscow State University, Faculty of Computational Mathematics and Cybernetics
Leninskie Gory 1, Bldg. 52, Moscow 119991, Russian Federation

⁴ Private healthcare institution Central Clinical Hospital "RZHD-Medicine"
Budayskaya Str. 2, Moscow 129128, Russian Federation

Background. Currently, machine learning and artificial intelligence technologies are rapidly developing and being implemented in scientific research, which significantly simplifies the study and use applied developments in medicine. The development of a method for automatic classification of Non-Hodgkin's B-cell lymphomas (NHL) and Chronic Lymphocytic Leukemia (CLL) in comparison with the norm (N) will allow us to analyze the differences in multidimensional relationships in the "immunity – tumor growth" system and will create the ability to classify these diseases based on blood data without using materials from other organs and tissues for diagnosis.

Purpose of this study – is to create an automated classification of NHL and CLL in comparison with the norm and to evaluate the quality of diagnostics based on differences in the indicators of multidimensional images of immunocompetent and tumor cells in peripheral blood.

Methodology. A comparative study of immunity and tumor cells in the peripheral blood of 185 healthy individuals, 352 patients with B-cell NHL, and 315 patients with CLL was conducted. The machine processing method classifying data by constructing a decision tree (DT) was used to compare the indicators of multidimensional images formed by lymphocyte subpopulations and tumor cells in the peripheral blood. Three classifiers of states were constructed: NHL versus normal (N–NHL), CLL versus normal (N–CLL), and NHL versus CLL (NHL–CLL). The quality of the classifiers was determined by the error's matrix and ROC curve criteria.

Results. The classifiers of values and distribution patterns of immunity and tumor cell indices constructed by the DT method differentiate between the states of N–NHL, N–CLL and NHL–CLL and determine the classification algorithm. The criteria of the error's matrix and ROC curve methods confirmed the high quality of the classifiers.

Conclusion. Using the machine learning method, a DT was constructed that classifies NHL and CLL, which allows for the implementation of alternative diagnostics of NHL and CLL using only data on the presence of immune and tumor lymphocytes in the peripheral blood.

Key words: peripheral blood; immunity; tumor lymphocytes; Non-Hodgkin's B-cell lymphomas; Chronic Lymphocytic Leukemia; machine classification; diagnostic algorithm.

For citation: Cherkashenco V.N., Saburina I.N., Kuzmina E.G., Antonov K.N., Ustinova E.E., Shchipkova E.S., Zacarenko S.V., Mushkarina T.Yu., Morozov S.G. [New approaches to the diagnosis of Non-Hodgkin's B-cell lymphomas and Chronic Lymphocytic Leukemia by clusters of differentiation of normal and tumor lymphocytes in peripheral blood]. *Patogenez [Pathogenesis]*. 2024; 22(4): 64–77. (in Russian)

DOI: 10.25557/2310-0435.2024.04.64-77

For correspondence: Cherkashenco Vladimir Nikolaevich, e-mail: comrade_cherkashenko@mail.ru

Funding. The study had no sponsorship.

Conflict of interests. The authors declare no conflict of interest.

Received: 05.12.2024.

Введение

Повышение продолжительности жизни часто связывают с ассоциированными с возрастом лимфопролиферативными заболеваниями – неходжкинскими В-клеточными лимфомами (НХЛ) и хроническим лимфолейкозом (ХЛЛ), актуализируя разработку экономически обоснованных быстрых методов диагностики этой патологии.

Для улучшения диагностики исследователи, как правило, отбирают диагностические и предсказательные факторы среди часто встречающихся признаков изучаемой патологии. При НХЛ и ХЛЛ к таким признакам следует отнести различия выхода опухолевых лимфоцитов в периферическую кровь. Поражение костного мозга и выход в кровь (лейкемизация) аномальных лимфоцитов свойственно ХЛЛ, при НХЛ наблюдается

значительно реже и менее выражено [1–3]. В патогенезе зрелоклеточной лимфопролиферации наблюдаются различные нарушения противоопухолевого иммунитета, приводящие к неэффективному взаимодействию иммунной системы с опухолью, к снижению, вплоть до полной утраты, контроля генетического гомеостаза организма [4, 5]. Однако использование отдельных различающихся признаков для диагностики болезни малоэффективно, т.к. они присутствуют не у каждого человека с диагностируемой патологией. Эффективность диагностики можно повысить путем интегрированного учёта признаков, различающих сравниваемые патологии. В наше время этого можно достигнуть, используя методы искусственного интеллекта (ИИ), в частности, машинного обучения (МО) [6, 7]. В обзорах [8–19] собраны недавно полученные результаты использования методов МО в решении различных медицинских диа-

гностических задач, подтверждающие значительный потенциал данных методов. В настоящее время методы МО демонстрируют многообещающие перспективы в анализе патологических изображений, обеспечивая решение трёх важнейших клинических задач: способствуют более точной диагностике; предоставляют информацию по прогнозу течения болезни в процессе лечения при установленном диагнозе; индивидуализируют варианты лечения. Использование методов МО для диагностики разных патологий сердечно-сосудистой системы, патологии выводящей системы почек, для рака молочной железы, диабета, болезни Паркинсона и болезни Альцгеймера обеспечивают точность диагностики от 70 до 98% [8–19].

В диагностике, лечении и прогнозировании НХЛ применяется широкий спектр методов МО [20]. Например, для решения диагностических задач НХЛ по гистологическим изображениям использован класс сверточных нейронных сетей [21]; по размерам увеличенных шейных лимфоузлов использована сеть глубокого обучения [22]; для дифференциальной диагностики лимфом, ассоциированных со слизистой оболочкой орбиты и конъюнктивы — метод машины опорных векторов [23]; для дифференциальной диагностики хронических лимфопролиферативных заболеваний по опухолевым и нормальным В-клеткам — метод случайного леса [24]; для модели прогноза диффузной В-клеточной лимфомы по ключевым генным маркерам генома — метод машины опорных векторов [25]; модель прогноза диффузной В-клеточной лимфомы от проведенной комбинированной химиотерапии построена на основе метода дерева решения [26].

В отечественной медицине активно применяются инструментальные методы обследования пациентов с НХЛ, в частности, проточная цитофлуориметрия с широким спектром клеточных маркеров. При этом, однако, не уделяется должного внимания методам МО для построения диагностических процедур. Так, в работе Волченко Н.Н. с соавторами методом проточной цитометрии, регистрирующей широкий спектр поверхностных клеточных маркеров (иммунофенотипирование), классифицированы хронические формы лейкозов (включая ХЛЛ) и лимфом (НХЛ) [27]. Диагностику проводили в два этапа, для каждого из которых были сформированы свои панели моноклональных антител. На первом этапе панель включала 15 маркеров, определяющих линейность опухолевых клеток (Т, В, НК). На втором этапе, в зависимости от линейности, панели содержали от 18 до 24 маркеров. При этом в работе не использованы методы МО. В работе Новицкого А.В. с соавт. решалась задача иммунодиагностики НХЛ. При использовании 26 моноклональных антител диагноз обосновывался качественными признаками. Во внимание не принималась многомерность анализируемых признаков [28].

Ранее нами было показано наличие чётких различающихся многомерных образов в анализируемых дан-

ных пациентов НХЛ и ХЛЛ между собой и в сравнении с нормой по количеству находящихся в периферической крови опухолевых и иммунных клеток [29, 30]. Это позволило предположить, что методы МО могут быть использованы для построения автоматического классификатора высокого качества для повышения эффективности диагностики заболеваний.

Цель — создать простую, максимально точную автоматизированную классификацию для диагностики НХЛ, ХЛЛ в сравнении с нормой по кластерам дифференцировки иммунокомпетентных и опухолевых клеток, находящихся в периферической крови, формирующих многомерные образы этих состояний.

Для достижения данной цели были поставлены следующие задачи:

1. На основе данных общепринятой диагностики (эталонной) состояния НХЛ, ХЛЛ и Нормы, разработать алгоритмы автоматизированной классификации, выявив методами машинной обработки различия многомерных структур по маркерам иммунокомпетентных и опухолевых клеток, находящихся в периферической крови.

2. Выявить значимые классифицирующие признаки и их иерархию в ходе автоматизированной классификации НХЛ, ХЛЛ методом дерева решений.

3. Оценить перспективу классификации для диагностики НХЛ, ХЛЛ, используя критерии качества методов матрицы ошибок и ROC кривой.

Материалы и методы исследования

Группы пациентов. Обследованы 352 пациента с разными вариантами В-клеточных НХЛ и 315 пациентов с типичными и атипичными формами ХЛЛ. Данные 185 доноров крови составили контрольную группу практически здоровых людей (норма).

Для диагностики ХЛЛ и НХЛ были взяты подходы стандартизации исследований, предложенные в рекомендациях ВОЗ [1–2, 27, 28] и принятые нами за эталон.

ХЛЛ представляет собой опухоль из В-лимфоцитов, первичная пролиферация и накопление которых происходят в костном мозге, лимфоидных органах и тканях с ранним выходом опухолевых клеток в кровь. При НХЛ лимфома растёт локально, преимущественно в лимфоузлах, иногда в тканях экстранодально, реже поражается костный мозг, в основном при прогрессировании заболевания происходит лейкемизация с выходом клеток в кровь.

Диагноз ХЛЛ и НХЛ установлен по критериям классификации ВОЗ опухолей лимфоидной и кроветворной системы, разработанных коллективами авторов, изложенных в международных и отечественных рекомендациях и статьях [1, 2]. Алгоритм диагностики предусматривает объективный учет данных ультразвукового, магнитно-резонансного обследования пациентов, иммуногистохимического и иммунофенотипического исследования материала лимфатических узлов/перифери-

ческой крови и костного мозга; определения зрелости и иммунофенотипа опухолевых В-лимфоцитов; учёт распространения (стадия заболевания по Rai и Binet) и молекулярно-генетических изменений. Первичная диагностика лимфопролиферативного заболевания (НХЛ и ХЛЛ) включает характеристику субстрата опухолевого роста, распространенность процесса (стадия заболевания), оценку общего состояния больного (прогностический индекс).

Как правило, для диагностики НХЛ и ХЛЛ применяют морфологическое исследование биопсийного или операционного материала. При первичном обследовании проводится гистологическое и иммуногистохимическое исследование инцизионной или эксцизионной биопсии патологического очага или операционного материала.

Метод проточной цитометрии позволяет идентифицировать по поверхностным молекулам опухолевые и нормальные клетки кроветворной и иммунной системы.

При наличии в общем анализе крови или в миелограмме лимфоцитоза преобладания лимфоидных клеточных элементов, атипичных лимфоцитов и подозрительных пролиферирующих клеток проводят иммунофенотипирование методами иммуногистохимии или проточной цитометрии для установления зрелости и характеристик лимфопролиферативного процесса.

Для анализа методом проточной цитофлуориметрии материалом служат клетки крови, костного мозга, выпотных жидкостей, бронхоальвеолярного смыва, ликвора, гомогенизированные образцы тканей (селезёнка, лимфоузлы и т.д.), клеточная суспензия, полученная при аспирационной тонкоигльной пункции лимфоузлов. Проточное иммунофенотипирование лимфоцитов выполняется с использованием 4–6 цветной панели моноклональных антител.

Диагностика различных типов хронических форм лимфопролиферативных заболеваний ХЛЛ и НХЛ проводится в два этапа, используя специально сформированные панели моноклональных антител. Сначала подобранными маркерами определяется линейность опухолевых клеток (Т-, В-, НК-лимфоцитов). Далее, в зависимости от линейности, определяется тип лимфопролиферативного заболевания. Диагностические критерии ХЛЛ: моноклональные лимфоциты >5000 /мкл крови; фенотип CD19⁺, CD5⁺, CD23⁺, CD79b⁺dim, CD20⁺dim, CD22⁺dim; рестрикция лёгких цепей (каппа либо лямбда); $>30\%$ лимфоцитов в костном мозге.

НХЛ диагностируется при наличии лимфаденопатии, и/или спленомегалии, и/или цитопении, вызванной поражением костного мозга, при условии, что абсолютное число моноклональных В-лимфоцитов в крови не превышает 5000/мкл.

Отметим, что до настоящего времени используемые подходы не принимают во внимание многомерность используемых параметров. Классификация проводится по каждому отдельному признаку.

Для создания процедуры автоматической диагностики использовали показатели крови, определяемые методами многопараметровой проточной цитометрии (цитофлуориметры FACScan или FACSCanto II, (BD Biosciences). Нами были отобраны те из них, которые формируют кластеры и клеточные структуры при изучаемой патологии [29, 30]. Это: абсолютное количество лейкоцитов, процентное количество лимфоцитов, Т-клеток (CD3, CD4, CD8, CD3HLADR, иммунорегуляторный индекс или соотношение CD4/CD8), количество в процентном отношении НК-лимфоцитов с фенотипом CD16⁺CD3⁻ и В-лимфоцитов (CD19, CD20, CD22, CD23). Был учтён также процент HLADR-лимфоцитов (сумма нормальных, опухолевых В-клеток и активированных Т-лимфоцитов). Для нормы учитывали В-клетки по экспрессии CD19, не оценивая CD20, CD22, CD23. В исследование были взяты данные пациентов, в банке данных которых были представлены полностью все 10(13) показателей.

Методы машинного обучения. Результаты машинного анализа признаков заболевания, представленные в виде дерева решений (ДР) (Decision Tree), часто используют для диагностики [31, 32]. Построение ДР относят к классу методов обучения с учителем. Поэтому классификация проводилась в два последовательных этапа: Этап построения классификатора на обучающей выборке и оценка качества полученного ДР на тестовой выборке с расчётом показателей качества. На первом этапе были использованы данные сравниваемых групп (НХЛ, ХЛЛ, норма), т.е. признаков, характеризующих сопоставляемые классифицируемые группы – маркёры иммунных и опухолевых клеток. На обучающей выборке были получены модели в виде ДР, определяющие правила интерпретации последовательно отобранных классифицирующих признаков, различающих сравниваемые заболевания и проведена классификация пар: НХЛ в сравнении с нормой (Н-НХЛ), ХЛЛ в сравнении с нормой (Н-ХЛЛ), НХЛ в сравнении с ХЛЛ (НХЛ-ХЛЛ). На втором этапе классифицированы тестовые данные с помощью построенного ДР. Для тестирования точности классификации были введены только показатели, характеризующие классифицируемые группы, без обозначения принадлежности к группам. Далее по результатам тестирования была оценена точность классификации.

Для построения ДР использована среда математического моделирования Matlab. Структура ДР представлена множеством узлов, соединённых ветвями. Узлы – логические выражения, ветви определяют переходы между ними (рис. 1).

Формирование решения начинается с корневого узла (вершина ДР), продолжается движением по ветвям ДР вправо или влево в зависимости от величины признака следующего узла. Каждый узел – правило, утверждающее значение кодируемого им признака. Проход по ДР отмечает значение переменных, характеризующих классифицируемые группы. Процесс про-

верки логических выражений завершается в терминальном узле [32].

Оценка качества классификации. Решение основной задачи классификации – возможности использования её для диагностики – базируется на оценке её качества [33–34]. В нашем исследовании для оценки качества классификаторов использованы два метода – матрица ошибок и ROC-кривая (рис. 2).

Как видно из рис. 2, матрица ошибок – это комбинация нашей классификации (классы «1» или «0» при классификации 2 классов) и фактического значения анализируемого объекта («1» или «0»). В зависимости

от результата предсказания и корректности проведения классификации, матрица разделена на 4 квадрата, в которых расположены абсолютные показатели качества классификации (TP, TN, FP и FN), которые используются для оценки качества классификационных моделей. Например, показатель true positive (TP, истинно положительный) результат – это количество случаев, в которых алгоритм правильно классифицирует входной признак как положительный. В свою очередь, показатель true negative (TN, истинно отрицательный) характеризует ситуацию, когда объект класса «0» классифицируется как относящийся именно к этому классу.

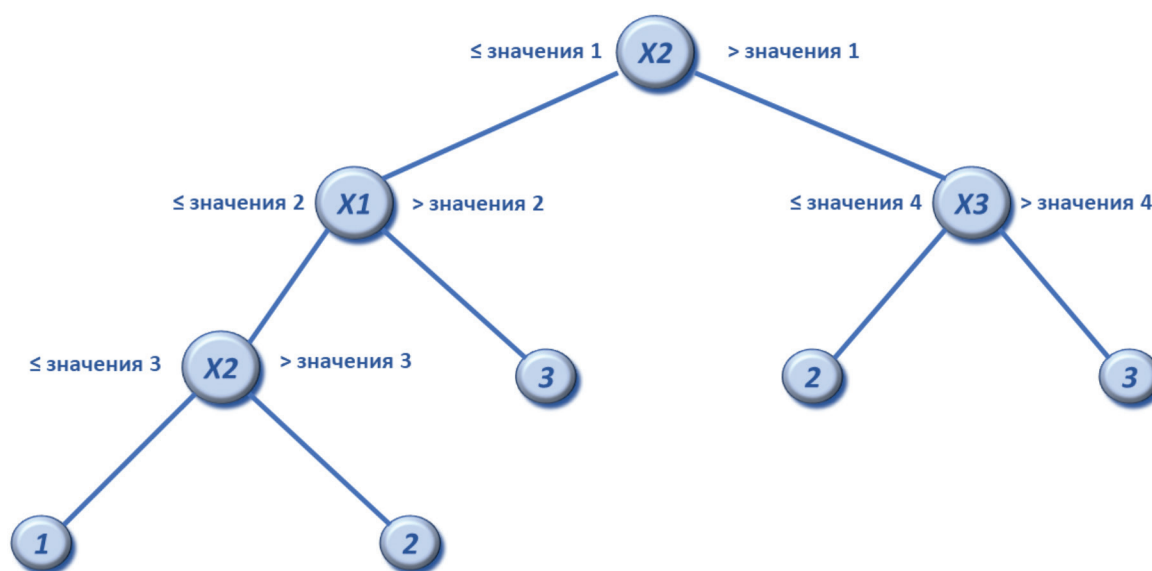


Рис. 1. Структура дерева решений для абстрактного классификатора.

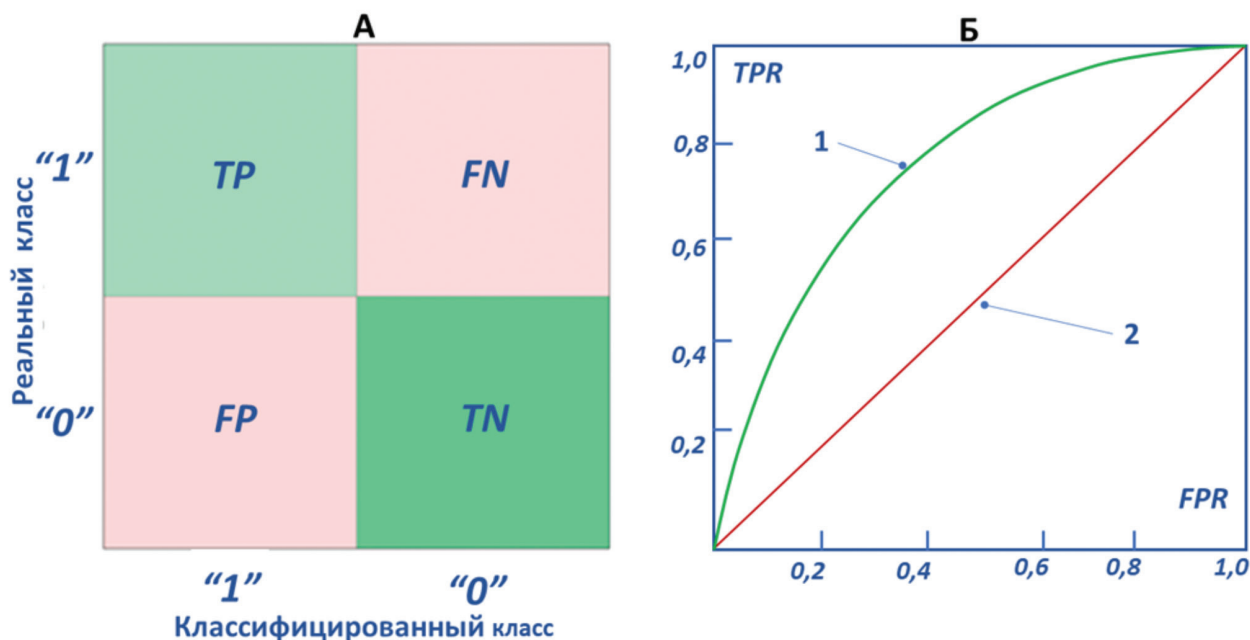


Рис. 2. Методы оценки качества классификационной модели: А – матрица ошибок; Б – ROC-кривая.

Как видно из рисунка, показатели TP и TN располагаются на главной диагонали матрицы ошибок. Соответственно, показатель false positive (FP, ложноположительный) определяет число случаев, в которых ошибочно классифицируется входной признак как положительный (класс «1») тогда, когда он принадлежит к другому классу (классу «0»). Показатель FN (False Negatives) – положительные объекты, классифицированные как отрицательные, то есть это событие есть «ложный пропуск» – событие ошибочно не обнаруживается (ложно отрицательные случаи). То есть, данный случай представляет собой «ложное обнаружение», так как при отсутствии события ошибочно выносится решение о его присутствии (ложно положительные случаи).

Эффективная классификация характеризуется матрицей ошибок, у которой на главной диагонали представлены большие значения, а на дополнительной диагонали – маленькие значения.

Для оценки качества классификации используется несколько производных показателей, основанных на компонентах матрицы ошибок (абсолютных оценках). Эти показатели являются относительными показателями, которые представлены в **табл. 1**.

ROC-кривая (**рис. 2, Б**) представляет ещё один инструмент оценки качества анализируемой модели. Это графическая характеристика качества бинарного классификатора – зависимость доли верных положительных классификаций от доли ложных положительных классификаций при варьировании уровня отсечения решающего правила, то есть, зависимость TPR от FPR.

Прямой 2 на **рис. 2** соответствует плохой классификатор с отсутствием классифицирующей способности, тогда как кривая 1 характеризует классификатор с имеющейся, пусть и незначительной, классифицирующей возможностью. ROC кривая позволяет отли-

чать плохой классификатор от хорошего; чем ближе она располагается к левому верхнему углу, тем лучше классификатор. Количественной характеристикой классификатора по ROC кривой является показатель AUC (Area Under Curve), равный площади под ROC-кривой. Возможные значения AUC (в %) варьируют в интервале [0; 100]. Приемлемыми считаются классификаторы с $AUC \geq 70\%$.

Результаты исследования и обсуждение

Алгоритм классификации, т.е. последовательность принятия решений выбора показателей, обеспечивающих различия групп Н-НХЛ и Н-ХЛЛ представлен на **рис. 3**.

Классификация Н-НХЛ. Основные признаки и их размерность, разделяющие норму и НХЛ, отражены в 16 вариантах ДР: 4 комбинации признаков классифицируют данные как норму, 12 – как НХЛ (**рис.3, А**).

Классификатор использовал 9 из 10 взятых в анализ признаков, что указывает на вовлеченность функционально различающихся типов лимфоцитов в патогенез НХЛ. Последовательным сравнением классификатор выделил из клеточного состава опухолевых и иммунных клеток в крови пограничные, разделительные значения показателей при дифференцировании нормы и НХЛ, и их иерархию. Основным корневым признаком дифференциации нормы и НХЛ классификатор использовал маркер активации – HLA-DR, который в норме локализуется на мембране активированных Т-лимфоцитов и нормальных В-клеток. При НХЛ этот маркер локализуется также и на опухолевых В-клетках.

В правой ветви принято 5 решений для НХЛ и 3 для нормы. Основные классифицирующие признаки: нарастание пограничных значений процен-

Таблица 1.

Метрики для оценки качества диагностики в методах машинного обучения.

№	Название оценки качества классификатора	Математическое выражение
1	Ложно положительный коэффициент (false positive rate) - FPR	$FPR = \frac{FP}{FP + TN}$
2	Ложно отрицательный коэффициент (false negative rate) - FNR	$FNR = \frac{FN}{FN + TP}$
3	Чувствительность (sensitivity или TPR)	$TPR = \frac{TP}{TP + FN}$
4	Специфичность (specificity или TPN)	$TPN = \frac{TN}{TN + FP}$
5	Точность (precision)	$Precision = \frac{TP}{TP + FP}$
6	Правильность (accuracy)	$Accuracy = \frac{TP + TN}{n}$
7	F1-score (гармоническое среднее)	$F1 = \frac{2 * Precision * Sensitivity}{Precision + Sensitivity}$

та HLA-DR и соотношения CD4/CD8, процентов В- и NK-клеток. Последовательность шагов принятия решений обусловлена степенью выраженности измененного признака, частотой его встречаемости в данной группе пациентов и сопряженностью с другими признаками. Различия сопоставляемых признаков определяют особенности геометрии многомерно-го образа нормы или НХЛ, а различия размерности этих отклонений указывают на особенности патогенеза общей группы НХЛ.

В левой ветви ДР выделено 1 правило для определения нормы и 7 правил для НХЛ. Поэтапно, для дифференцировки классификатор подключил более низкий при НХЛ по сравнению с нормой уровень в крови CD19В-клеток. Далее, ДР последовательно относит данные в группу НХЛ, ориентируясь на различия в размерности процента лимфоцитов, общей популяции CD3Т-клеток, CD8Т-киллеров, CD4/CD8 и количества лейкоцитов. И, в заключение, выделяет практически классическую размерность формулы

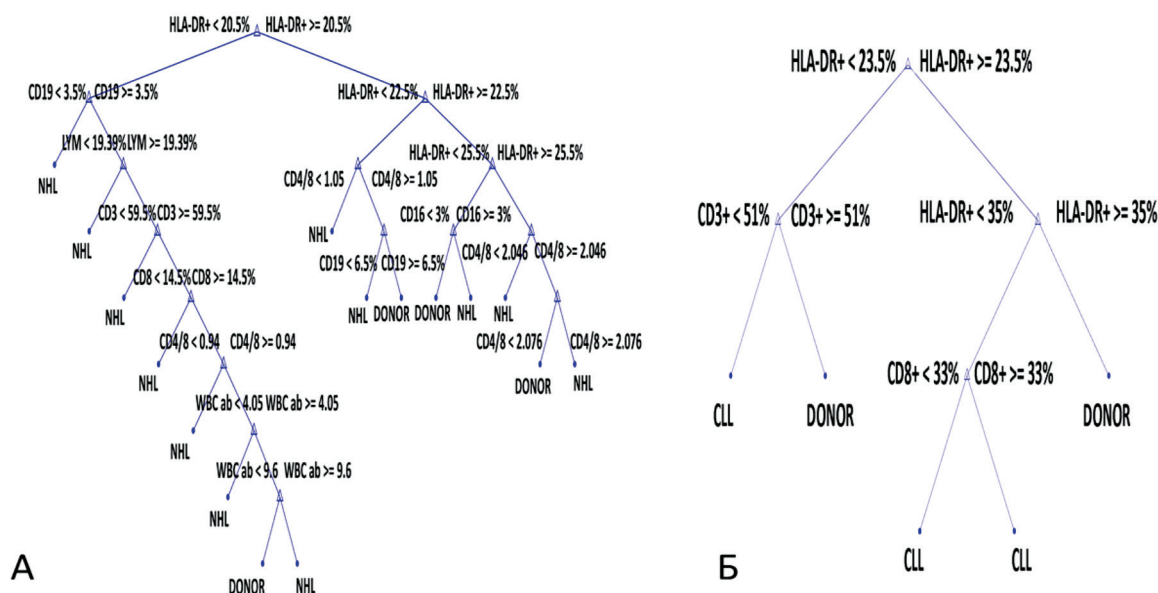


Рис. 3. Деревья решений для классификаторов патологических состояний: А – для Н-НХЛ; Б – для Н-ХЛЛ.

Таблица 2.

Классификационные признаки ДР для нормы и НХЛ

№ п/п	НОРМА	НХЛ
1	Лц>4,05, но <9,6 × 10 ⁹ кл/л, Лф>19,4%; CD3>59,5%; CD8>14,5%; CD4/CD8>0,96; CD19>3,5%; HLA-DR<20,5%.	HLA-DR<20,5%; CD19<3,5%
2	HLA-DR>20,5%, но <22,5%; CD4/CD8>1,06; CD19>6,5%	HLA-DR<20,5%; CD19>3,5%; Лф<19,39%
3	HLA-DR>22,5, но <25,5%, CD16<3,0%	HLA-DR<20,5%; CD19>3,5%; Лф>19,39% CD3<59,5%
4	CD4/CD8>2,04, но<2,076; HLA-DR>25,5%	HLA-DR<20,5%; CD19>3,5%; Лф>19,39%; CD3>59,5%; CD8<14,5%
5		HLA-DR<20,5%; CD19>3,5%; Лф>19,39%; CD3>59,5%; CD8>14,5%; CD4/CD8<0,96
6		HLA-DR<20,5%; CD19>3,5%; Лф>19,39%; CD3>59,5%; CD8>14,5%; CD4/CD8>0,96; Лц<4,05 × 10 ⁹ кл./л
7		HLA-DR<20,5%; CD19>3,5%; Лф>19,3%; CD3> 59,5%; CD8>14,5%; CD4/CD8>0,96; Лц>9,6 × 10 ⁹ кл./л
8		HLA-DR> 20,5%, но < 22,5; CD4/CD8<1,05
9		HLA-DR>20,5%, но < 22,5%; CD4/CD8>1,05; CD19< 6,5%
10		HLA-DR>25,5%; CD16>3,0%
11		HLA-DR>25,5%; CD4/CD8<2,045
12		HLA-DR>25,5%; CD4/CD8>2,076

иммунофенотипа Нормы, описанную во многих руководствах: Лц $>4,05$, но $<9,6 \times 10^9$ кл./л, Лф $>19,4\%$; CD3 $>59,5\%$; CD8 $>14,5\%$; CD4/CD8 $>0,96$; CD19 $>3,5\%$, HLA-DR $<20,5\%$.

В табл. 2 приведены правила, на основании которых построенная модель проводит классификацию значений и выносит диагностическое суждение для нормы и НХЛ.

Классификация Н-ХЛЛ. Путь ДР Н-ХЛЛ содержит 3 уровня иерархии (против 8 у Н-НХЛ) и включает 5 правил (рис. 3, Б). Классификатор структуры образцов для дифференцировки ХЛЛ и нормы в качестве основных различающих признаков выделяет 3 показателя крови: проценты HLA-DR лимфоцитов, общих CD3Т-клеток и CD8Т-киллеров. Патологическая составляющая этой триады при ХЛЛ общеизвестна. При ХЛЛ в крови значительно повышен процент В-клеток за счет роста жизнеспособных опухолевых В-лимфоцитов со сниженным уровнем апоптоза; также повышен уровень активированных Т-клеток, экспрессирующих маркер активации – CD3HLA-DR. Классификатор выделяет клетки суммарно по общему маркеру HLA-DR, указывая на связь в патогенезе ХЛЛ динамики прироста объема опухолевых клеток с уровнем активации иммунитета. Более низкое значение второго узлового признака, процента общих CD3Т-клеток при НХЛ, обусловлено преобладанием выхода и задержкой в крови опухолевых В-клеток по сравнению со снижением Т-клеточного иммунного ответа по мере развития и прогрессирования ХЛЛ. От-

бор процента CD8Т-киллеров в качестве третьего дифференцирующего признака свидетельствует о повышении специфической и неспецифической функциональной нагрузки на Т-киллеры при ХЛЛ, реагирующие на антигены инфицированных, опухолевых и мутировавших клеток. Феномен истощения CD8Т-киллеров и его связь с неблагоприятным течением заболевания отмечен в исследовании [35].

Классификатор выделяет 2 сочетания признаков для отнесения анализируемых данных к норме и 3 сочетания к ХЛЛ (табл. 3).

Классификация НХЛ–ХЛЛ. Для построения ДР при классификации «НХЛ–ХЛЛ» использованы 10 показателей, как и при классификации НХЛ и ХЛЛ с нормой. В ДР для классификатора, различающего ХЛЛ от НХЛ, количество уровней иерархии равно 5, правил решения 11 (рис. 4). Признаки, выделенные классификатором для принятия решений об отнесении данных к НХЛ или ХЛЛ, те же, что при дифференцировке НХЛ и ХЛЛ в сравнении с нормой, но иерархия и величина пограничных значений различаются. При обеих формах заболевания, как ХЛЛ, так и НХЛ, дифференцирующим показателем, вынесенным в корневой узел, является превышающее верхнюю границу нормы количество лейкоцитов, указывающее, что оба заболевания развиваются на фоне воспалительных процессов разной степени выраженности с усилением активации противомикробного и противоопухолевого иммунитета.

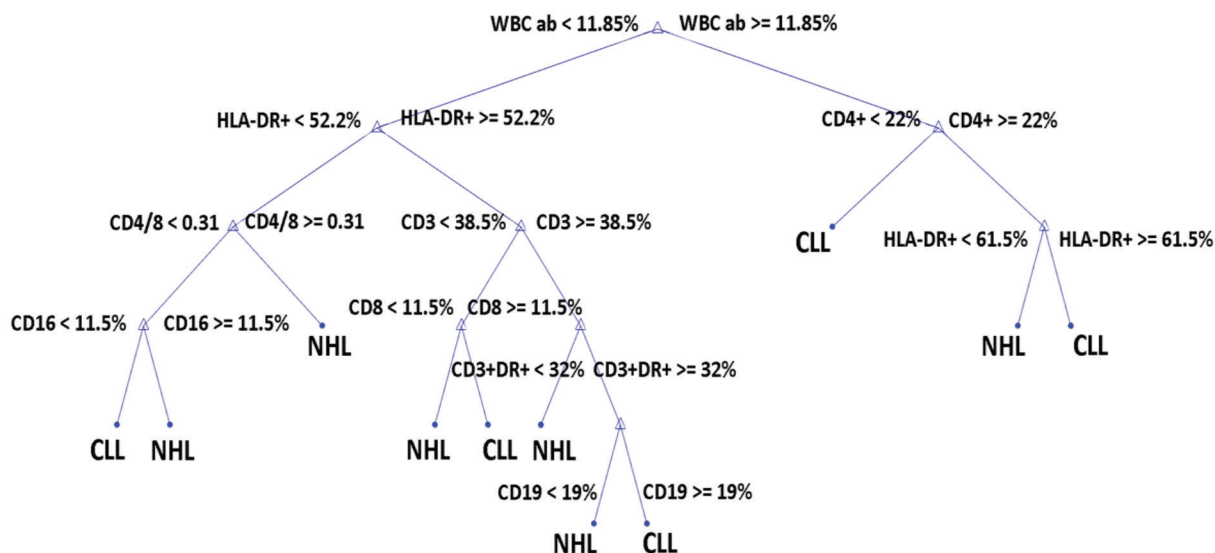


Рис. 4. Дерево решений для классификатора заболеваний ХЛЛ – НХЛ.

Таблица 3.

Классификационные признаки ДР для нормы и ХЛЛ

№ п/п	НОРМА	ХЛЛ
1	HLA-DR $<23,5\%$; CD3 $>51\%$	HLA-DR $<23,5\%$ при CD3 $<51\%$
2	HLA-DR $>23,5$, но $<35,5\%$; CD8 $<33\%$	HLA-DR $>35,5\%$
3		HLA-DR $>23,5\%$, но $<35,5\%$, CD8 $>33\%$

Вторые узлы ДР представлены количеством CD4⁺ Т-хелперов и HLA-DR-клеток в процентном соотношении 22 и 52 соответственно.

Третий уровень иерархии определяется различием CD3⁺ Т-клеток (38,5%), соотношением CD4/CD8 Т-лимфоцитов (0,31%) в левой ветви ДР и уровнем HLA-DR активированных лимфоцитов (62,2%) в правой ветви. На четвертой ступени добавляются различия в уровне активации CD3⁺HLA-DR⁺ Т-клеток (32%), NK-клеток (11,5%) и CD8 Т-киллеров (11,5%). На пятой ступени различия затрагивают CD19 В-клетки.

Участие в патогенезе 9 из 10 взятых в анализ признаков указывает на определенные различия вовлеченности разных типов лимфоцитов в патогенез НХЛ и ХЛЛ. Классификатор в приоритетные различающие признаки этих заболеваний выносит на первый план количество лейкоцитов, интегрально отражающее процессы воспаления и лейкоемизации, по-разному проявляющиеся в крови при НХЛ и ХЛЛ. Признаки активации иммунитета, эффективность которых ограничивается недостаточностью CD4 Т-хелперов, являются определяющими эти процессы. При этом, по данным крови, количество киллеров врожденного иммунитета, NK-клеток, и специфического иммунитета, CD8 Т-киллеров критически различаются при НХЛ и ХЛЛ. Именно эти различия использованы в качестве критерия диагностики этих заболеваний.

Классификатор дифференцировал пространственные образы НХЛ и ХЛЛ. Получены 5 решений о диагнозе ХЛЛ и 6 решений о диагнозе НХЛ (табл. 4)

Оценка качества моделей. Как уже было отмечено, классификаторы могут быть применены для диагностики заболевания в зависимости от качества классификации признаков, т.е. вероятности, отражающей встречаемость выявляемых различий. Оценка качества построенных моделей классификации представлена на графиках ROC кривыми. (рис. 5).

На рис. 5 представлены ROC-кривые: в левой части для классификатора Н-НХЛ, (рис. 5, А); на среднем графике – для классификатора Н-ХЛЛ (рис. 5, Б); и в правой части – для классификатора ХЛЛ-НХЛ (рис. 5, В).

Все три ROC-кривые проходят очень близко к верхнему левому углу и имеют высокие показатели AUC. Самый хороший классификатор по показателю AUC – это классификатор Н-ХЛЛ, AUC = 0,98. Для распознавания Н-НХЛ -AUC = 0,92, а для дифференциальной диагностики НХЛ-ХЛЛ этот показатель равен 0,89.

Все три модели с высокой эффективностью классифицируют состояния Н-НХЛ, Н-ХЛЛ и НХЛ-ХЛЛ.

Но, как известно, AUC не является единственным показателем качества. Другие показатели характеризуют разные аспекты работы классификаторов. Причем, все они определены на отрезке [0, 1], и чем ближе их значение к 1, тем выше качество классификатора по данному показателю.

На рис. 6 по метрикам качества (табл. 1) построены диаграммы классификаторов Н-НХЛ, Н-ХЛЛ и НХЛ-ХЛЛ.

Анализ показывает, что наилучшее качество по пяти критериям дает классификатор Н-ХЛЛ (рис. 6, Б). Все пять метрик качества имеют близкие высокие значения; четыре из них выше 0,95 (кроме показателя «специфичность»). Для классификатора Н-НХЛ метрики качества наиболее низкие в группах сравнения (рис. 6, А), но достаточно высокие, чтобы пользоваться данным классификатором. Такие метрики, как «точность» и «специфичность» выше, чем у классификатора НХЛ-ХЛЛ (рис. 6, В). Этот классификатор занимает промежуточное положение между Н-НХЛ и Н-ХЛЛ. Значения его метрик на уровне 0,9.

Таким образом, выполненная классификация с высокой эффективностью различает многомерные образы, формируемые лейкоцитами и опухолевыми клетками в крови, и может быть использована для их диагностики.

Заключение

Сравнительный анализ различных сочетаний находящихся в периферической крови опухолевых клеток и различающихся вариантов нарушений специфического и неспецифического иммунитета позволил разработать новые подходы к диагностике хронических

Таблица 4.

Классификационные признаки ДР для диагностики ХЛЛ и НХЛ.

№ п/п	ХЛЛ	НХЛ
1	Лц > 11,85 × 10 ⁹ кл./л; CD4 < 22%	Лц > 11,85 × 10 ⁹ кл./л; CD4 > 22,0%; HLA-DR < 61,5%
2	Лц > 11,85 × 10 ⁹ кл./л; CD4 > 22%; HLA-DR > 61,5%.	Лц < 11,85 × 10 ⁹ кл./л; HLA-DR < 52,2%; CD4/CD8 > 0,31
3	Лц < 11,85 × 10 ⁹ кл./л; HLA-DR < 52,2%; CD4/CD8 < 0,31; CD16 < 11,5%	Лц < 11,85 × 10 ⁹ кл./л; HLA-DR < 52,2%; CD4/CD8 < 0,31; CD16 > 11,5%
4	Лц < 11,85 × 10 ⁹ кл./л; HLA-DR > 52,2%; CD3 < 38,5%; CD8 > 11,5%.	Лц < 11,85 × 10 ⁹ кл./л; HLA-DR > 52,2%; CD3 < 38,5%; CD8 < 11,5%
5	Лц < 11,85 × 10 ⁹ кл./л; HLA-DR > 52,2%; CD3 > 38,5%; CD3HLA-DR > 32,0%; CD19 > 19,0%	Лц < 11,85 × 10 ⁹ кл./л; HLA-DR > 52,2%; CD3 > 38,5%; CD3HLA-DR < 32,0%.
6		11,85 × 10 ⁹ кл./л; HLA-DR > 52,2%; CD3 > 38,5%; CD3HLA-DR > 32,0%; CD19 < 19,0%

лимфопролиферативных заболеваний – ХЛЛ и В-клеточных НХЛ.

Современными методами машинной обработки данных (построение ДР) проведено многомерное сравнение и создан алгоритм автоматизированной классификации ХЛЛ, НХЛ в сравнении с нормой. Построены классификаторы, которые различают в крови многомерные структуры состояний Н–НХЛ, Н–ХЛЛ и ХЛЛ–ХЛЛ и протестирован их алгоритм. Проверка качества классификаторов методами матрицы ошибок и ROC-кривых показали высокое качество анализа (превышают 0,90) по чувствительности методов для всех групп сравне-

ния. Они позволяют определять состояние НХЛ и ХЛЛ в сравнении с нормой, используя данные о содержании в крови количества лейкоцитов, лимфоцитов, Т-лимфоцитов (CD3Т-общих, CD4Т-хелперов CD8Т-киллеров, их соотношения, CD4/CD8, активированных CD3HLA-DRT-клеток), CD16NK-клеток, нормальных и опухолевых В-лимфоцитов (CD19, CD20, CD22, CD23), и HLA-DR-лимфоцитов. Результаты исследования могут быть использованы для альтернативной диагностики НХЛ и ХЛЛ, а также предполагают возможность дальнейшей разработки классификаторов для диагностики подтипов НХЛ.

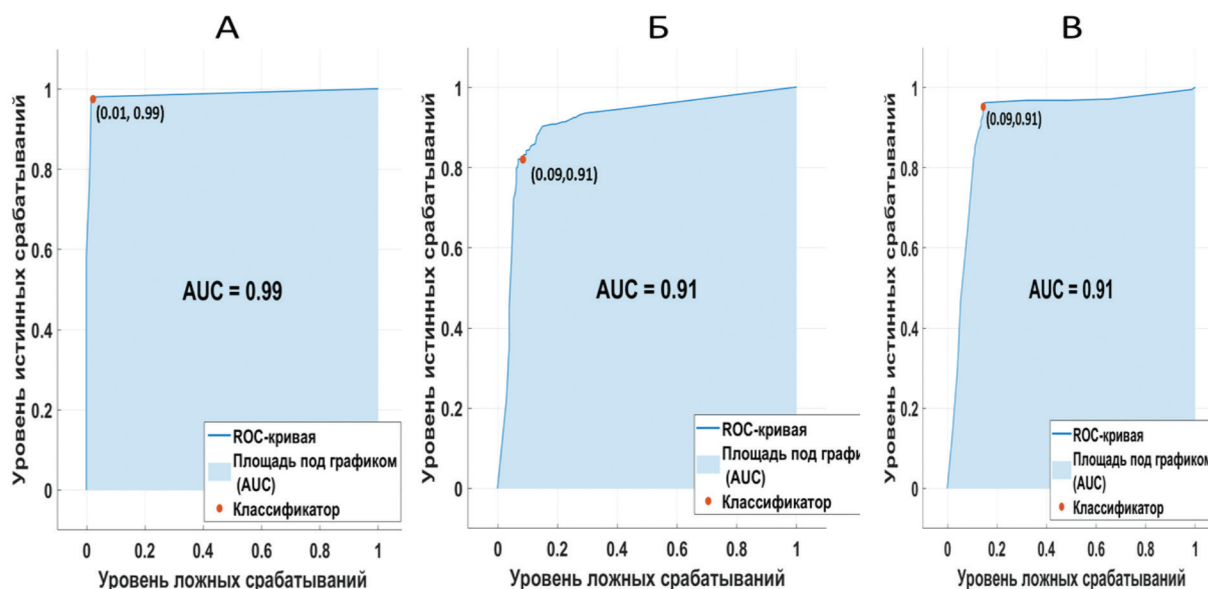


Рис. 5. ROC-кривые для трех классификаторов: А – для различения Н–НХЛ, Б – для различения Н–ХЛЛ, В – для различения ХЛЛ–НХЛ.

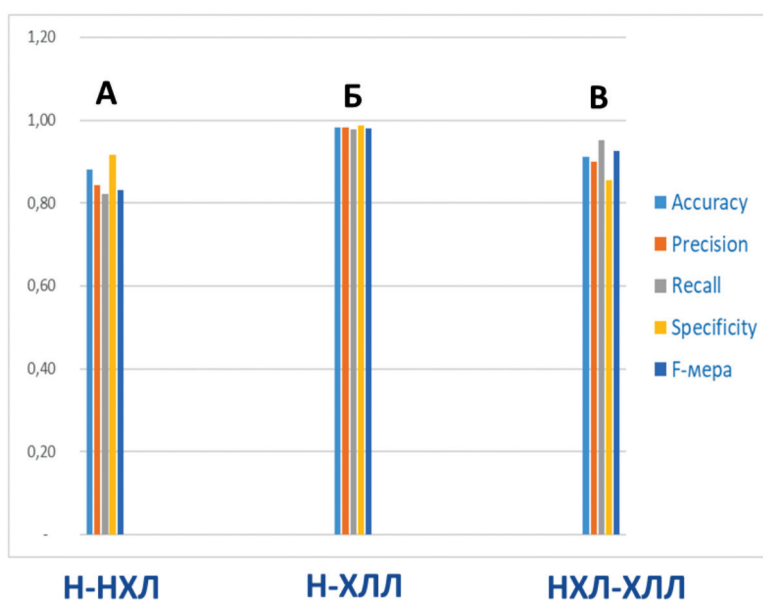


Рис. 6. Метрики качества классификаторов: А – Н- НХЛ, Б – Н-ХЛЛ, В – ХЛЛ-НХЛ.

Авторский вклад

Черкашенко В.Н. — концепция, выбор и применение методов машинного обучения, подготовка рукописи и иллюстративного материала, редактирование текста статьи; Сабурова И.Н. — концепция, подготовка и редактирование рукописи, общее руководство исследованием; Кузьмина Е.Г. — анализ данных проточной цитофлуориметрии, создание базы клинических и иммунологических данных, клиничко-иммунологическая интерпретация данных полученных классификаторов, подготовка рукописи, редактирование текста статьи; Антонов К.Н. — программирование используемых алгоритмов, численные эксперименты с используемыми данными, получение результатов моделирования; Устинова Е.Е. и Щепина Е.С. — подготовка иллюстративного материала; Зацаренко С.В., Мушкарина Т.Ю. — иммунофенотипирование; Меркулова Д.М., Морозов С.Г. — редактирование текста статьи; окончательное одобрение рукописи — все авторы.

Список литературы

1. Хронический лимфолейкоз. Современная диагностика и лечение. Под ред. Османова Е.А., Румянцева А.Г. Москва: «ГЭОТАР-Медиа», 2023. 480 с. DOI: 10.33029/9704-7597-3-SDL-2023-1-480/ – EDN OXZFQT
2. Campo E., Jaffe E.S., Cook J.R., Quintanilla-Martinez L., Swerdlow S.H., Anderson K.C., Brousset P., Cerroni L., de Leval L., Dimhofer S., Dogan A., Feldman A.L., Fend F., Friedberg J.W., Gaulard Ph., Ghia P., Horwitz S.M., King R.L., Salles G., San-Miguel J., Seymour J.F., Treon S.P., Vose J.M., Zucca E., Advani R., Ansell S., Au W.-Y., Barrionuevo C., Bergsagel L., Chan W.C., Cohen J.I., d'Amore F., Davies A., Falini B., Ghobrial I.M., Goodlad J.R., Gribben J.G., Hsi E.D., Kahl B.S., Kim W.-S., Kumar S., LaCasce A.S., Laurent C., Lenz G., Leonard J.P., Link M.P., Lopez-Guillermo A., Mateos M.V., Macintyre E., Melnick A.M., Morschhauser F., Nakamura S., Narbaitz M., Pavlovsky A., Pileri S.A., Piris M., Pro B., Rajkumar V., Rosen S.T., Sander B., Sehn L., Shipp M.A., Smith S.M., Staudt L.M., Thieblemont C., Tousseyn T., Wilson W.H., Yoshino T., Zinzani P.-L., Dreyling M., Scott D.W., Winter J.N., Zelenetz A.D. The International Consensus Classification of Mature Lymphoid Neoplasms: a report from the Clinical Advisory Committee. *Blood*. 2022; 140(11): 1229–1253. DOI: 10.1182/blood.2022015851
3. Неприна Г.С., Кузьмина Е.Г., Павлов В.В., Шахтарина С.В., Двинских Н.Ю., Рогова Н.М., Константинова Т.В., Курасова В.Г., Сироткина Н.П. Иммунофенотипический анализ лимфоцитов костного мозга и периферической крови у больных В-клеточными злокачественными лимфомами. *Медицинская иммунология*. 2006; 8(2–3): 347–348. DOI: 10.15789/1563-0625-2006-2-3-329-360
4. Mitta I.D., Gubin M.M., Schreiber R.D., Smyth M.J. New insights into cancer immunoeediting and its three component phases — elimination, equilibrium and escape. *Curr Opin. Immunol.* 2014; 27: 16–25. DOI: 10.1016/j.coi.2014.01.004
5. Hacken E., Burger J.A. Microenvironment interactions and B cell receptor signaling in chronic lymphocytic leukemia: implications for disease pathogenesis and treatment. *Biochim. Biophys. Acta*. 2016; 1863(3): 401–413. DOI: 10.1016/j.bbamcr.2015.07.009
6. Caballé-Cervigón N., Castillo-Sequera J.L., Gomez-Pulido J.A., Gómez J.M. Machine Learning Applied to Diagnosis of Human Diseases: A Systematic Review. *Appl. Sci.* 2020; 10(15): 5135. DOI: 10.3390/app10155135
7. Ibrahim M.I., Adnan M.A. The Role of Machine Learning Algorithms for Diagnosing Diseases. *Journal of applied science and technology trends*. 2021; 2(1): 10–19. DOI: 10.38094/jastt20179
8. Singh H.N.V., Navaneeth G.N. Pillai et al *Multisurface Proximal SVM Based Decision Trees For Heart Disease Classification*. In Proceedings of the TENCON 2019 IEEE Region 10 Conference (TENCON), Kerala, India, 17–20 October 2019; 13–18. DOI: 10.1109/TENCON.2019.8929618
9. Desai S.D., Giraddi S., Narayankar P., Pudakalakatti N.R., Sulegaon Sh. *Back-propagation neural network versus logistic regression in heart disease classification*. In *Advanced Computing and Communication Technologies*. Springer: Berlin/Heidelberg, Germany, 2019: 133–144.
10. Kumar R.R., Polepaka S. *Performance Comparison of Random Forest Classifier and Convolution Neural Network in Predicting Heart Diseases*. In ICCII 2018, Proceedings of the Third International Conference on Computational Intelligence and Informatics. Springer: Singapore, 2020: 683–691. DOI: 10.1007/978-981-15-1480-7_59
11. Nithya A., Appathurai A., Venkatadri N., Ramji D.R. Kidney disease detection and segmentation using artificial neural network and multi-kernel k-means clustering for ultrasound images. *Measurement*. 2019; 149: 106952. DOI: 10.1016/j.measurement.2019.106952
12. Brunetti A., Cascarano G., De Feudis I., Moschetta M. *Detection and segmentation of kidneys from magnetic resonance images in patients with autosomal dominant polycystic kidney disease*. In Proceedings of the International Conference on Intelligent Computing. Nanchang, China, 3–6 August 2019; Springer: Berlin/Heidelberg, Germany, 2019: 639–650. DOI: 10.1007/978-3-030-26969-2_60
13. Moon W.K., Chen I.-L., Chang J.M., Shin S.U., Lo Ch.-M., Chang R.-F. The adaptive computer-aided diagnosis system based on tumor sizes for the classification of breast tumors detected at screening ultrasound. *Ultrasonics*. 2017; 76: 70–77. DOI: 10.1016/j.ultras.2016.12.017
14. Abdel-Nasser M., Domenec H.R., Moreno P.A. Analysis of tissue abnormality and breast density in mammographic images using a uniform local directional pattern. *Expert Syst. Appl.* 2015; 42(24): 9499–9511. DOI: 10.1016/j.eswa.2015.07.072
15. Qtea H., Awad M. Using Hybrid Model of Particle Swarm Optimization and Multi-Layer Perceptron Neural Networks for Classification of Diabete. *Int. J. Intell. Eng. Syst.* 2021; 14(3): 11–22. DOI: 10.22266/ijies2021.0630.02
16. Fitriyani N.L., Ganjar M.S., Rhee A.J. Development of disease prediction model based on ensemble learning approach for diabetes and hypertension. *IEEE Access*. 2019; 7: 144777–144789. DOI: 10.1109/ACCESS.2019.2945129
17. Shetty S., Rao Y. *SVM based machine learning approach to identify Parkinson's disease using gait analysis*. In Proceedings of the 2016 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 26–27 August 2016; IEEE: Manhattan, NY, USA, 2016; 2: 1–5. DOI: 10.1109/INVENTIVE.2016.7824836
18. Ouhmida A., Raihani A., Cherradi B., Sara S. *Parkinson's disease identification using KNN and ANN Algorithms based on Voice Disorder*. In Proceedings of the 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco, 19–20 March 2020; IEEE: Manhattan, NY, USA, 2020: 1–6. DOI: 10.1109/IRASET48871.2020.9092228
19. Sun M. *Automatic Diagnosis of Alzheimer's Disease and Mild Cognitive Impairment Based on CNN+ SVM Networks with End-to-end Training*. In Proceedings of the 2021 13th International Conference on Advanced Computational Intelligence (ICACI), Wanzhou, Chongqing, China, 14–16 May 2021; IEEE: Manhattan, NY, USA, 2021: 279–285.
20. Yuan J., Zhang Y., Wang X. Application of machine learning in the management of lymphoma: Current practice and future prospects, Review. *DIGITAL HEALTH*. 2024; 10: 1–12. DOI: 10.1177/20552076241247963
21. Steinbuss G., Kriegsmann M., Zgorzelski Ch., Brobeil A., Goepfert B., Dietrich S., Mechttersheimer G., Kriegsmann K. Deep learning for the classification of non-Hodgkin lymphoma on histopathological images. *Cancers (Basel)*. 2021; 13(10): 2419. DOI: 10.3390/cancers13102419
22. Can S., Türk Ö., Ayral M., Kozan G., Arı H., Akdağ M., Baylan M.Y. Can deep learning replace histopathological examinations in the differential diagnosis of cervical lymphadenopathy? *Eur. Arch. Otorhinolaryngol.* 2024; 281(1): 359–367. DOI: 10.1007/s00405-023-08181-9
23. Tagami M., Nishio M., Katsuyama-Yoshikawa A., Misawa N., Sakai A., Haruna Y., Azumi A., Hondaet Sh. Machine learning model with texture analysis for automatic classification of histopathological images of ocular adnexal mucosa-associated lymphoid tissue lymphoma of two different origins. *Curr. Eye Res.* 2023; 48(12): 1195–1202. DOI: 10.1080/02713683.2023.2246696
24. Gross Z., Veyrat-Masson R., Grange B., Huet S., Verney A., Traverse-Glehen A., Ruminy Ph., Baseggioet L. Diagnosis of chronic B-cell lymphoproliferative disease in peripheral blood = how machine learning may help to the interpretation of flow cytometry data.

- Hematol. Oncol.* 2024; 42(1): e3245. DOI: 10.1002/hon.3245
25. Sun J.B., Ke X.Y., Zhang M.Z., Wang Y., An F., Zhao Y., Zhuo L. New prognostic models for extranodal natural killer T-cell lymphoma, nasal-type using Cox regression and machine learning. *Transl. Cancer Res.* 2021; 10(2): 613–626. DOI: 10.21037/tcr-20-3017
 26. Zhao K., Zheng X., Liu X., Liu T., Ke Z., Zhu F., Wen Q., Xin B., Li Q., Zhang L. Tissue-matched IgH gene rearrangement of circulating tumor DNA shows significant value in predicting the progression of diffuse large B cell lymphoma. *Oncologist.* 2024; 29(5): 672–680. DOI: 10.1093/oncolo/oyae008
 27. Волченко Н.Н., Славнова Е.Н., Наумова Е.В. Роль проточной цитофлюориметрии в цитологической диагностике неходжкинских лимфом. *Клиническая лабораторная диагностика.* 2012; 8: 42–48.
 28. Новицкий А.В., Никитин В.Ю., Сухина И.А., Тыренко В.В., Иванов А.М. Проточная цитометрия в дифференциальной диагностике хронических В-клеточных лимфолипролиферативных заболеваний. *Вестник Санкт-Петербургской медицинской академии последипломного образования.* 2011; 3(2): 149–165.
 29. Кузьмина Е.Г., Черкашенко В.Н., Павлов В.В., Мушкарина Т.Ю., Антонов К.Н., Устинова Е.Е., Богданенко Е.В., Сабурин И.Н. Анализ структурных особенностей многомерных данных периферической крови при хроническом лимфоцитарном лейкозе и неходжкинской В-лимфоме в сравнении с нормой. *Патологическая физиология и экспериментальная терапия.* 2024; 68(1): 55–68. DOI: 10.25557/0031-2991.2024.01.55-6829
 30. Сабурин И.Н., Сабурин И.Н., Кузьмина Е.Г., Черкашенко В.Н., Павлов В.В., Зацаренко С.В., Антонов К.Н., Богданенко Е.В., Устинова Е.Е., Морозов С.Г. Сравнительный анализ структурных особенностей многомерных данных периферической крови при хроническом лимфоцитарном лейкозе и неходжкинской В-лимфоме. *Патологическая физиология и экспериментальная терапия.* 2024; 68(1): 69–80. DOI: 10.25557/0031-2991.2024.01.69-8030
 31. Кафтанников И.Л., Парасич А.В. Особенности применения деревьев решений в задачах классификации. *Вестник Южно-Уральского государственного университета. Серия «Компьютерные технологии, управление, радиоэлектроника».* 2015; 15(3): 26–32. DOI: 10.14529/ctcr150304
 32. Rokach L., Maimon O. *Decision Trees* in book: *The Data Mining and Knowledge Discovery.* Handbook, 2005: 165–192.
 33. Cali C., Longobardi M. Some mathematical properties of the ROC curve and their applications. *Ricerche di Matematica. CORE.* 2015; 64(2): 8. DOI: 10.1007/s11587-015-0246-8
 34. Мухамадиев Р.И., Мухамедиева Е.Л., Кучин Я.И. Таксономия методов машинного обучения и оценка качества классификации и обучаемости. *Электронный журнал Cloud Science.* 2015; 2(3): 359–378.
 35. Hanna B.S., Roessner P.M., Yazdanparast H., Colomer D., Campo E., Kugler S., Yosifov D., Stilgenbauer S., Schmidt M., Gabriel R., Lichter P., Seiffert M. Control of chronic lymphocytic leukemia development by clonally-expanded CD8⁺ T-cells that undergo functional exhaustion in secondary lymphoid tissues. *Leukemia.* 2019; 33(3): 625–637. DOI: 10.1038/s41375-018-0250-6
- of Mature Lymphoid Neoplasms: a report from the Clinical Advisory Committee. *Blood.* 2022; 140(11): 1229–1253. DOI: 10.1182/blood.2022015851
3. Neprina G.S., Kuzmina E.G., Pavlov V.V., Shakhtarina S.V., Dvinskikh N.Yu., Rogova N.M., Konstantinova T.V., Kurasova V.G., Sirotkina N.P. [Immunophenotypic analysis of bone marrow lymphocytes and peripheral blood in patients B-cell malignant lymphomas]. *Meditsinskaya immunologiya [Medical Immunology].* 2006; 8(2–3): 347–348. DOI: 10.15789/1563-0625-2006-2-3-329-360 (in Russian)
 4. Mitta I.D., Gubin M.M., Schreiber R.D., Smyth M.J. New insights into cancer immunoediting and its three component phases – elimination, equilibrium and escape. *Curr Opin. Immunol.* 2014; 27: 16–25. DOI: 10.1016/j.coi.2014.01.004
 5. Hacken E., Burger J.A. Microenvironment interactions and B cell receptor signaling in chronic lymphocytic leukemia: implications for disease pathogenesis and treatment. *Biochim. Biophys. Acta.* 2016; 1863(3): 401–413. DOI: 10.1016/j.bbamer.2015.07.009
 6. Caballé-Cervigón N., Castillo-Sequera J.L., Gomez-Pulido J.A., Gómez J.M. Machine Learning Applied to Diagnosis of Human Diseases: A Systematic Review. *Appl. Sci.* 2020; 10(15): 5135. DOI: 10.3390/app10155135
 7. Ibrahim M.I., Adnan M.A. The Role of Machine Learning Algorithms for Diagnosing Diseases. *Journal of applied science and technology trends.* 2021; 2(1): 10–19. DOI: 10.38094/jastt20179
 8. Singh H.N.V., Navaneeth G.N. Pillaiet al *Multisurface Proximal SVM Based Decision Trees For Heart Disease Classification.* In Proceedings of the TENCON 2019 IEEE Region 10 Conference (TENCON), Kerala, India, 17–20 October 2019; 13–18. DOI: 10.1109/TENCON.2019.8929618
 9. Desai S.D., Giraddi S., Narayankar P., Pudakalakatti N.R., Sulegaon Sh. *Back-propagation neural network versus logistic regression in heart disease classification.* In *Advanced Computing and Communication Technologies.* Springer: Berlin/Heidelberg, Germany, 2019: 133–144.
 10. Kumar R.R., Polepaka S. *Performance Comparison of Random Forest Classifier and Convolution Neural Network in Predicting Heart Diseases.* In ICCII 2018, Proceedings of the Third International Conference on Computational Intelligence and Informatics. Springer: Singapore, 2020: 683–691. DOI: 10.1007/978-981-15-1480-7_59
 11. Nithya A., Appathurai A., Venkatadri N., Ramji D.R. Kidney disease detection and segmentation using artificial neural network and multi-kernel k-means clustering for ultrasound images. *Measurement.* 2019; 149: 106952. DOI: 10.1016/j.measurement.2019.106952
 12. Brunetti A., Cascarano G., De Feudis I., Moschetta M. *Detection and segmentation of kidneys from magnetic resonance images in patients with autosomal dominant polycystic kidney disease.* In Proceedings of the International Conference on Intelligent Computing. Nanchang, China, 3–6 August 2019; Springer: Berlin/Heidelberg, Germany, 2019: 639–650. DOI: 10.1007/978-3-030-26969-2_60
 13. Moon W.K., Chen I-L., Chang J.M., Shin S.U., Lo Ch-M., Chang R-F. The adaptive computer-aided diagnosis system based on tumor sizes for the classification of breast tumors detected at screening ultrasound. *Ultrasonics.* 2017; 76: 70–77. DOI: 10.1016/j.ultras.2016.12.017
 14. Abdel-Nasser M., Domenec H.R., Moreno P.A. Analysis of tissue abnormality and breast density in mammographic images using a uniform local directional pattern. *Expert Syst. Appl.* 2015; 42(24): 9499–9511. DOI: 10.1016/j.eswa.2015.07.072
 15. Qtea H., Awad M. Using Hybrid Model of Particle Swarm Optimization and Multi-Layer Perceptron Neural Networks for Classification of Diabete. *Int. J. Intell. Eng. Syst.* 2021; 14(3): 11–22. DOI: 10.22266/ijies2021.0630.02
 16. Fitriyani N.L., Ganjar M.S., Rhee A.J. Development of disease prediction model based on ensemble learning approach for diabetes and hypertension. *IEEE Access.* 2019; 7: 144777–144789. DOI: 10.1109/ACCESS.2019.2945129
 17. Shetty S., Rao Y. *SVM based machine learning approach to identify Parkinson's disease using gait analysis.* In Proceedings of the 2016 International Conference on Invention Computation Technologies (ICICT), Coimbatore, India, 26–27 August 2016; IEEE: Manhattan, NY, USA, 2016; 2: 1–5. DOI: 10.1109/INVENTIVE.2016.7824836
 18. Ouhmida A., Raihani A., Cherradi B., Sara S. *Parkinson's disease identification using KNN and ANN Algorithms based on Voice Disorder.* In Proceedings of the 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco, 19–20 March 2020; IEEE: Manhattan, NY, USA, 2020: 1–6. DOI: 10.1109/IRASET48871.2020.9092228
 19. Sun M. *Automatic Diagnosis of Alzheimer's Disease and Mild Cognitive*

References

1. *Chronic lymphocytic leukemia. Modern diagnostics and treatment.* Ed.: Osmanova E.A., Rummyantseva A.G. Moscow: GEOTAR-Media, 2023. p. 480 DOI: 10.33029/9704-7597-3-SDL-2023-1-480/ – EDN OXZFQT (in Russian)
2. Campo E., Jaffe E.S., Cook J.R., Quintanilla-Martinez L., Swerdlow S.H., Anderson K.C., Brousset P., Cerroni L., de Leval L., Dirnhofer S., Dogan A., Feldman A.L., Fend F., Friedberg J.W., Gaulard Ph., Ghia P., Horwitz S.M., King R.L., Salles G., San-Miguel J., Seymour J.F., Treon S.P., Vose J.M., Zucca E., Advani R., Ansell S., Au W-Y., Barrionuevo C., Bergsagel L., Chan W.C., Cohen J.I., d'Amore F., Davies A., Falini B., Ghobrial I.M., Goodlad J.R., Gribben J.G., Hsi E.D., Kahl B.S., Kim W-S., Kumar S., LaCasce A.S., Laurent C., Lenz G., Leonard J.P., Link M.P., Lopez-Guillermo A., Mateos M.V., Macintyre E., Melnick A.M., Morschhauser F., Nakamura S., Narbaitz M., Pavlovsky A., Pileri S.A., Piris M., Pro B., Rajkumar V., Rosen S.T., Sander B., Sehn L., Shipp M.A., Smith S.M., Staudt L.M., Thieblemont C., Tousseyn T., Wilson W.H., Yoshino T., Zinzani P-L., Dreyling M., Scott D.W., Winter J.N., Zelenetz A.D. The International Consensus Classification

- Impairment Based on CNN+ SVM Networks with End-to-end Training*. In Proceedings of the 2021 13th International Conference on Advanced Computational Intelligence (ICACI), Wanzhou, Chongqing, China, 14–16 May 2021; IEEE: Manhattan, NY, USA. 2021: 279–285.
20. Yuan J., Zhang Y., Wang X. Application of machine learning in the management of lymphoma: Current practice and future prospects, Review. *DIGITAL HEALTH*. 2024; 10: 1–12. DOI: 10.1177/20552076241247963
 21. Steinbuss G., Kriegsmann M., Zgorzelski Ch., Brobeil A., Goepfert B., Dietrich S., Mechttersheimer G., Kriegsmann K. Deep learning for the classification of non-Hodgkin lymphoma on histopathological images. *Cancers (Basel)*. 2021; 13(10): 2419. DOI: 10.3390/cancers13102419
 22. Can S., Türk Ö., Ayril M., Kozan G., Ari H., Akdağ M., Baylan M.Y. Can deep learning replace histopathological examinations in the differential diagnosis of cervical lymphadenopathy? *Eur. Arch. Otorhinolaryngol.* 2024; 281(1): 359–367. DOI: 10.1007/s00405-023-08181-9
 23. Tagami M., Nishio M., Katsuyama-Yoshikawa A., Misawa N., Sakai A., Haruna Y., Azumi A., Hondaet Sh. Machine learning model with texture analysis for automatic classification of histopathological images of ocular adnexal mucosa-associated lymphoid tissue lymphoma of two different origins. *Curr. Eye Res.* 2023; 48(12): 1195–1202. DOI: 10.1080/02713683.2023.2246696
 24. Gross Z., Veyrat-Masson R., Grange B., Huet S., Verney A., Traverse-Glehen A., Ruminy Ph., Baseggioet L. Diagnosis of chronic B-cell lymphoproliferative disease in peripheral blood = how machine learning may help to the interpretation of flow cytometry data. *Hematol. Oncol.* 2024; 42(1): e3245. DOI: 10.1002/hon.3245
 25. Sun J.B., Ke X.Y., Zhang M.Z., WangY., An F., Zhao Y., Zhuet L. New prognostic models for extranodal natural killer T-cell lymphoma, nasal-type using Cox regression and machine learning. *Transl. Cancer Res.* 2021; 10(2): 613–626. DOI: 10.21037/tcr-20-3017
 26. Zhao K., Zheng X., Liu X., Liu T., Ke Z., Zhu F., Wen Q., Xin B., Li Q., Zhang L. Tissue-matched IgH gene rearrangement of circulating tumor DNA shows significant value in predicting the progression of diffuse large B cell lymphoma. *Oncologist.* 2024; 29(5): 672–680. DOI: 10.1093/oncolo/oyae008
 27. Volchenko N.N., Slavnova E.N. Naumova E.V. [The role of flow cytometry in the cytological diagnosis of non-Hodgkin's lymphomas]. *Klinicheskaya laboratornaya diagnostika [Scientific and Practical Journal]*. 2012; 8: 42–48. (in Russian)
 28. Novitsky A.V., Nikitin V.Yu., Sukhina I.A., Tyrenko V.V., Ivanov A.M. [Flow cytometry in the differential diagnosis of chronic B-cell lymphoproliferative diseases]. *Vestnik Sankt-Peterburgskoi medicinskoj akademii posleddiplomnogo obrazovaniya [Bulletin of the St. Petersburg Medical Academy of Postgraduate Education]*. 2011; 3(2): 149–165. (in Russian)
 29. Kuzmina E.G., Cherkashenko V.N., Pavlov V.V., Mushkarina T.Yu., Antonov K.N., Ustinova E.E., Bogdanenko E.V., Saburina I.N. [Analysis of structural features of multidimensional data of peripheral blood in chronic lymphocytic leukemia and non-Hodgkin's B-lymphoma in comparison with the norm]. *Patologicheskaya fiziologiya i eksperimental'naya terapiya [Pathological Physiology and Experimental Therapy]*. 2024; 68(1): 55–68. DOI: 10.25557/0031-2991.2024.01.55-6829 (in Russian)
 30. Saburina I.N., Saburina I.N., Kuzmina E.G., Cherkashenko V.N., Pavlov V.V., Zatsarenko S.V., Antonov K.N., Bogdanenko E.V., Ustinova E.E., Morozov S.G. [Comparative analysis of structural features of multidimensional data of peripheral blood in chronic lymphocytic leukemia and non-Hodgkin's B-lymphoma]. *Patologicheskaya fiziologiya i eksperimental'naya terapiya [Pathological Physiology and Experimental Therapy]*. 2024; 68(1): 69–80. DOI: 10.25557/0031-2991.2024.01.69-8030 (in Russian)
 31. Kaftannikov I.L., Parasich A.V. [Features of the application of decision trees in classification problems]. *Vestnik Yuzhno-Ural'skogo gosudarstvennogo universiteta. Seriya «Komp'yuternyye tekhnologii, upravleniye, radioelektronika» [Bulletin of the South Ural State University. Series "Computer technologies, control, radio electronics"]*. 2015; 15(3): 26–32. DOI: 10.14529/ctcr150304 31 (in Russian)
 32. Rokach L., Maimon O. *Decision Trees* in book: *The Data Mining and Knowledge Discovery*. Handbook, 2005: 165–192.
 33. Cali C., Longobardi M. Some mathematical properties of the ROC curve and their applications. *Ricerche di Matematica. CORE*. 2015; 64(2): 8. DOI: 10.1007/s11587-015-0246-8
 34. Mukhamadiev R.I., Mukhamedieva E.L., Kuchin Ya.I. [Taxonomy of machine learning methods and assessment of the quality of classification and learning ability]. *Elektronnyy zhurnal Cloud Science [Electronic journal Cloud Science]*. 2015; 2(3): 359–378. (in Russian)
 35. Hanna B.S., Roessner P.M., Yazdanparast H., Colomer D., Campo E., Kugler S., Yosifov D., Stilgenbauer S., Schmidt M., Gabriel R., Lichter P., Seiffert M. Control of chronic lymphocytic leukemia development by clonally-expanded CD8⁺ T-cells that undergo functional exhaustion in secondary lymphoid tissues. *Leukemia*. 2019; 33(3): 625–637. DOI: 10.1038/s41375-018-0250-6

Сведения об авторах:

Черкашенко Владимир Николаевич — научный сотрудник лаборатории клеточной биологии и

патологии развития Федерального государственного бюджетного научного учреждения «Научно-исследовательский институт общей патологии и патофизиологии»; <https://orcid.org/0009-0009-0071-471X>

Сабурина Ирина Николаевна — доктор биологических наук, главный научный сотрудник лаборатории клеточной биологии и патологии развития Федерального государственного бюджетного научного учреждения «Научно-исследовательский институт общей патологии и патофизиологии»; <https://orcid.org/0000-0003-2014-2535>

Кузьмина Евгения Геннадьевна — кандидат биологических наук, ведущий научный сотрудник отделения клинической иммунологии Медицинского радиологического научного центра имени А.Ф. Цыба — филиала Федерального государственного бюджетного учреждения «Научный медицинский исследовательский центр радиологии» Министерства здравоохранения Российской Федерации; <http://orcid.org/0000-0002-0728-593X>

Антонов Кирилл Николаевич — студент факультета вычислительной математики и кибернетики Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет имени М.В. Ломоносова»; <https://orcid.org/0009-0009-0091-3591>

Устинова Екатерина Евгеньевна — младший научный сотрудник лаборатории клеточной биологии и патологии развития Федерального государственного бюджетного научного учреждения «Научно-исследовательский институт общей патологии и патофизиологии»; <https://orcid.org/0009-0004-6238-9100>

Щипкова Екатерина Сергеевна — младший научный сотрудник лаборатории клеточной биологии и патологии развития Федерального государственного бюджетного научного учреждения «Научно-исследовательский институт общей патологии и патофизиологии»; <https://orcid.org/0009-0008-8976-0212>

Зацаренко Светлана Валерьевна — научный сотрудник отделения клинической иммунологии Медицинского радиологического научного центра имени А.Ф. Цыба — филиала Федерального государственного бюд-

жетного учреждения «Научный медицинский исследовательский центр радиологии» Министерства здравоохранения Российской Федерации; <https://orcid.org/0000-0002-6335-1401>

Мушкарina Татьяна Юрьевна — научный сотрудник отделения клинической иммунологии Медицинского радиологического научного центра имени А.Ф. Цыба — филиала Федерального государственного бюджетного учреждения «Научный медицинский исследовательский центр радиологии» Министерства здравоохранения Российской Федерации; <https://orcid.org/0000-0002-1266-1792>

Меркулова Дина Мироновна — доктор медицинских наук, профессор, руководитель неврологического центра имени профессора Б.М. Гехта Частного учреждения здравоохранения Центральная клиническая больница «РЖД-Медицина»; <https://orcid.org/0000-0003-0368-683X>

Морозов Сергей Георгиевич — доктор медицинских наук, профессор, член-корреспондент РАН, директор Федерального государственного бюджетного научного учреждения «Научно-исследовательский институт общей патологии и патофизиологии»; <https://orcid.org/0000-0001-5822-5729>