

DOI: <https://doi.org/10.17816/EID601806>

Применение алгоритма дерева решений для ранней дифференциальной диагностики между различными клиническими формами острого иксодового клещевого боррелиоза и клещевого энцефалита

Е.Н. Ильинских, Е.Н. Филатова, К.В. Самойлов, А.В. Семёнова, С.В. Аксёнов

Сибирский государственный медицинский университет, Томск, Российская Федерация

АННОТАЦИЯ

Обоснование. Клещевой энцефалит и иксодовый клещевой боррелиоз являются наиболее распространёнными в России природно-очаговыми инфекциями, нередко протекающими в виде смешанной формы, которую зачастую клинически сложно отличить от моноинфекции в начале болезни, что может быть обусловлено поздней лабораторной верификацией диагноза и требует поиска принципиально нового подхода к проблеме раннего дифференциального диагноза клещевых инфекций.

Цель исследования — создание алгоритмов деревьев решений для ранней дифференциальной диагностики между изолированными и смешанными формами острого иксодового клещевого боррелиоза и клещевого энцефалита с преобладанием лихорадочного синдрома на основе анализа клинико-лабораторных данных.

Материалы и методы. Ретроспективно проанализированы 55 клинико-лабораторных параметров, полученных в первую неделю болезни от 291 госпитализированного больного клещевыми инфекциями, имевшего или не имевшего мигрирующую эритему на месте присасывания иксодового клеща, которые были включены в одноцентровое исследование в период с 2010 по 2023 год. У 211 больных без эритемы анализ проводился между тремя классами в зависимости от диагноза: имевшими смешанную инфекцию безэритемной формы иксодового клещевого боррелиоза с клещевым энцефалитом, моноинфекцию безэритемного боррелиоза или моноинфекцию клещевого энцефалита. Два других класса, включавших 80 пациентов с эритемой, имели микст-заболевание острого эритемного боррелиоза с клещевым энцефалитом или боррелиозную моноинфекцию. Язык программирования Python был применён для разработки двух моделей деревьев решений. Определялись показатели важности всех предикторов. Каждый из классов пациентов был случайным образом разделён на обучающую (70%) и тестовую (30%) выборки. Оценка точности моделей была основана на ROC-анализе.

Результаты. Алгоритм дерева решений для ранней дифференциальной диагностики у больных клещевыми инфекциями без эритемы включал следующие наиболее важные предикторы: максимальную высоту лихорадки, озноб, индекс соотношения нейтрофилов и моноцитов, скорость оседания эритроцитов, абсолютное число реактивных лимфоцитов и незрелых гранулоцитов, а также процентное содержание эозинофилов. Модель для дифференциального диагноза между пациентами с эритемой включала: высоту лихорадки, абсолютное число реактивных лимфоцитов и незрелых гранулоцитов, а также процентное содержание базофилов. Обе модели деревьев решений получили высокую прогностическую оценку на основании определения чувствительности, специфичности, прецизионности, точности и показателя F1, а также площади под ROC-кривой, превышающей 0,90.

Заключение. С использованием клинико-лабораторных параметров разработаны два алгоритма дерева решений, имеющие высокую чувствительность, которые легко применимы в клинической практике для ранней дифференциальной диагностики клещевых инфекций с преобладанием синдрома лихорадки.

Ключевые слова: дерево решений; машинное обучение; иксодовый клещевой боррелиоз; клещевой энцефалит; дифференциальный диагноз.

Как цитировать

Ильинских Е.Н., Филатова Е.Н., Самойлов К.В., Семёнова А.В., Аксёнов С.В. Применение алгоритма дерева решений для ранней дифференциальной диагностики между различными клиническими формами острого иксодового клещевого боррелиоза и клещевого энцефалита // Эпидемиология и инфекционные болезни. 2023. Т. 28, № 5. С. 275–288. DOI: <https://doi.org/10.17816/EID601806>

Рукопись получена: 01.10.2023

Рукопись одобрена: 17.10.2023

Опубликована: 24.10.2023

DOI: <https://doi.org/10.17816/EID601806>

Applying decision tree algorithms to early differential diagnosis between different clinical forms of acute Lyme borreliosis and tick-borne encephalitis

Ekaterina N. Ilyinskikh, Evgenia N. Filatova, Kirill V. Samoylov, Alina V. Semenova, Sergey V. Axyonov

Siberian State Medical University, Tomsk, Russian Federation

ABSTRACT

BACKGROUND: Tick-borne encephalitis and Lyme borreliosis are the most common natural focal infections in Russia, often arising as a mixed infection. Distinguishing these infections from mono-infections at the onset of the disease is often clinically challenging due to delayed laboratory verification of the diagnosis. It requires further searching for fundamentally new approaches for early differential diagnosis of tick-borne infections.

AIM: This study aimed to develop decision tree algorithms for early differential diagnosis between the mono- and mixed forms of acute Lyme borreliosis and tick-borne encephalitis with prevailing febrile syndrome in clinical picture based on clinical and laboratory data.

MATERIALS AND METHODS: We retrospectively analyzed 55 clinical and laboratory parameters obtained from 291 hospitalized patients with tick-borne infections, both with or without erythema migrans at the site of ixodid tick bites, during the first week of the disease. These patients were included in a single-center study conducted from 2010 to 2023. Among the 211 patients without erythema, the analysis was carried out between three classes depending on the diagnosis: mixed infection of nonerythematous Lyme borreliosis and tick-borne encephalitis, mono-infection of nonerythematous Lyme borreliosis, or mono-infection of tick-borne encephalitis. The other two classes that included 80 patients with erythema, had the mixed infection of acute erythematous Lyme borreliosis and tick-borne encephalitis or erythematous Lyme borreliosis mono-infection. Python programming language was used to develop two decision tree models. Feature importance was assessed for all predictors. Each patient class was randomly divided into training (70%) and testing (30%) datasets. Accuracy evaluation of the models was based on ROC analysis.

RESULTS: The decision tree algorithm for early differential diagnosis among patients with tick-borne infection without erythema migrans included most important predictors: maximal fever rise, chills, neutrophil-to-monocyte ratio, ESR, absolute number of reactive lymphocytes and immature granulocytes, and percentage of eosinophils. The model for differential diagnosis between the patients with erythema migrans included the following predictors: maximal fever rise, the absolute number of reactive lymphocytes and immature granulocytes, and the percentage of basophils. Both decision tree models demonstrated excellent predictive values based on sensitivity, specificity, precision, accuracy, and F1 scores, as well as areas under the ROC curve exceeding 0.90.

CONCLUSIONS: Based on clinical and laboratory parameters, two decision tree models with high sensitivity have been developed. These models can be easily applied in clinical practice for early differential diagnosis of the tick-borne infections with prevailing fever syndrome.

Keywords: decision tree; machine learning; Lyme borreliosis; tick-borne encephalitis; differential diagnosis.

To cite this article

Ilyinskikh EN, Filatova EN, Samoylov KV, Semenova AV, Axyonov SV. Applying decision tree algorithms to early differential diagnosis between different clinical forms of acute Lyme borreliosis and tick-borne encephalitis. *Epidemiology and Infectious Diseases*. 2023;28(5):275–288. DOI: <https://doi.org/10.17816/EID601806>

Received: 01.10.2023

Accepted: 17.10.2023

Published: 24.10.2023

ОБОСНОВАНИЕ

Среди инфекций, передающихся иксодовыми клещами, клещевой энцефалит (КЭ) и иксодовый клещевой боррелиоз (ИКБ) являются самыми распространёнными в России природно-очаговыми заболеваниями, которые нередко протекают в виде смешанной формы [1–3]. Эти заболевания имеют значительное сходство в клинической картине, в особенности на начальной стадии болезни, протекающей преимущественно с преобладанием лихорадочно-интоксикационного синдрома [4–6]. Поэтому нередко запаздывающая лабораторная верификация диагноза клещевых инфекций, обусловленная поздним появлением специфических противоборрелиозных антител, может приводить к назначению неадекватной этиотропной терапии, что потенциально способствует прогрессированию инфекционного процесса с развитием клинически более тяжёлых и/или хронических форм заболевания [7].

На сегодняшний день отсутствует простой инструмент, позволяющий дифференцировать изолированные и смешанные формы ИКБ и КЭ в первые дни болезни на основе клинических данных до получения результатов специфических лабораторных тестов. Методы машинного обучения в настоящее время используются для прогноза исхода и ранней дифференциальной диагностики инфекционных заболеваний, таких как лихорадка денге [8, 9] и коронавирусная инфекция COVID-19 [10, 11], в частности, это модели деревьев решений, представляющие собой эффективные алгоритмы классификации, которые выявляют нелинейные связи между предикторами, к основным преимуществам которых можно отнести наглядность и простоту интерпретации при практическом применении [12].

Цель исследования — разработка алгоритмов деревьев решений для ранней дифференциальной диагностики между изолированными и смешанными формами острого ИКБ и КЭ с преобладанием в клинической картине лихорадочно-интоксикационного синдрома на основе анализа клинико-лабораторных данных в начале заболевания.

МЕТОДЫ

Дизайн исследования

Одноцентровое ретроспективное выборочное наблюдательное исследование включало 291 больного со смешанной инфекцией (СИ) острого ИКБ и лихорадочной формы (ЛФ) КЭ или с моноинфекциями (МИ) этих заболеваний, которые соответствовали критериям включения.

Критерии соответствия

Критерии включения: больные клещевыми инфекциями в возрасте от 20 до 75 лет, госпитализированные не позднее 7 дней от начала заболевания с клиническими проявлениями преимущественно лихорадочно-интоксикационного

синдрома, которые имели клинико-эпидемиологическое и лабораторное подтверждение диагнозов СИ или МИ острого ИКБ и/или ЛФ КЭ.

Критериями исключения были беременность, лактация, менингеальный синдром, подтверждённый характерными лабораторными изменениями спинномозговой жидкости, энцефалит, другие острые и хронические инфекционные болезни (туберкулёз, хронические вирусные гепатиты В и С, ВИЧ-инфекция и др.), а также онкологическая или тяжёлая соматическая патология.

Условия проведения

Исследование проведено на базе инфекционной клиники Сибирского государственного медицинского университета (ФГБОУ ВО СибГМУ Минздрава России), г. Томск.

Продолжительность исследования

В исследование включены больные клещевыми инфекциями, госпитализированные в весенне-летний период эпидемических сезонов с 2010 по 2023 год.

Исходы исследования

Проанализированы клинические и лабораторные данные, полученные однократно при госпитализации из историй болезней пациентов, поступивших в стационар в первые 7 дней от начала болезни, с диагнозами СИ или МИ острого ИКБ и/или ЛФ КЭ. Кроме того, для верификации окончательного диагноза специфические лабораторные тесты были проведены в парных пробах сыворотки в день госпитализации пациента, а также в динамике через 14, 21 день, 3 и 6 месяцев. У обследованных нами больных были верифицированы 5 вариантов окончательных диагнозов, которые для построения дифференциально-диагностических моделей деревьев решений с использованием различных параметров клинико-лабораторных данных были разделены на две категории в зависимости от наличия или отсутствия в месте присасывания иксодового клеща мигрирующей эритемы — патогномичного признака ИКБ.

С использованием метода машинного обучения построены две модели деревьев решений, позволяющих проводить дифференциальный диагноз в первую неделю болезни до получения результатов специфических лабораторных тестов среди больных без эритемы на месте присасывания клеща, имевших три варианта окончательных диагнозов: СИ безэритемной формы (БФ) острого ИКБ с ЛФ КЭ, МИ БФ ИКБ или МИ ЛФ КЭ, а также среди больных с эритемой, имевших два варианта диагнозов: СИ эритемной формы (ЭФ) ИКБ с ЛФ КЭ или МИ ЭФ ИКБ.

Анализ в подгруппах

Для построения моделей деревьев решений каждый из исходных 5 классов пациентов с различными

окончательными диагнозами случайным образом были дополнительно разделены на две выборки — обучающую и тестовую с сохранением отношения 70 и 30% [12]. В дальнейшем для построения алгоритмов деревьев решений использовали только обучающие выборки, в то время как тестовые выборки применялись для валидации полученных алгоритмов и исключения переобучения, то есть ситуации, когда модель хорошо описывает данные из обучающей выборки, но плохо применима для тестовых данных.

Методы регистрации исходов

Окончательный диагноз у всех больных был поставлен на основании клинико-эпидемиологических данных и результатов специфических лабораторных тестов, оценка которых проводилась опытными клиницистами в динамике в период госпитализации, а также после выписки из стационара. Диагнозы изолированной и смешанной инфекции ИКБ и/или ЛФ КЭ были сформулированы в соответствии с их клиническими классификациями [1–3].

Для лабораторного подтверждения диагнозов ИКБ и/или КЭ применялся метод иммуноферментного анализа (ИФА) с определением в диагностических титрах специфических иммуноглобулинов (Ig) классов М и G к *Borrelia burgdorferi* s. l., а также антигена вируса КЭ, IgM и IgG к вирусу КЭ, используя тест-системы АО «Вектор-Бест» (Россия). Помимо этого, с помощью наборов серии «Реал-Бест» (АО «Вектор-Бест», Россия) для полимеразной цепной реакции (ПЦР) у пациентов были исключены гранулоцитарный анаплазмоз человека и эрлихиозы (*Anaplasma phagocytophilum*, *Ehrlichia muris*, *Ehrlichia chaffeensis*), а также возвратная клещевая лихорадка, вызываемая *Borrelia miyamotoi*.

В исследовании были проанализированы 29 клинических параметров, в том числе максимальная высота температуры тела и другие проявления лихорадочно-интоксикационного синдрома в первую неделю болезни. Кроме того, анализ включал 20 показателей общего и биохимического анализов крови при поступлении пациентов в стационар, в том числе стандартные и расширенные показатели гемограммы, такие как IG (immature granulocytes) — абсолютное количество незрелых гранулоцитов и RE-LYMP (reactive lymphocytes) — абсолютное количество реактивных лимфоцитов, определенных с помощью автоматического гематологического анализатора «Sysmex XN-1000» («Sysmex Corp.», Япония), а также 6 лейкоцитарных индексов интоксикации (ЛИИ), включая индекс соотношения нейтрофилов и моноцитов (ИСНМ), индекс соотношения лимфоцитов и моноцитов (ИСЛМ), нейтрофильно-лимфоцитарный индекс (НЛИ), индекс соотношения лейкоцитов и скорости оседания эритроцитов (ИСЛСОЭ), лимфоцитарно-гранулоцитарный индекс (ИЛГ) и ЛИИ по В.К. Островскому (ЛИИО) [13, 14].

Этическая экспертиза

От всех пациентов в письменном виде было получено информированное согласие. Проведение исследования было одобрено Локальным этическим комитетом ФГБОУ ВО СибГМУ Минздрава России (протоколы № 7939 от 21.10.2019 г. и № 9119/1 от 30.05.2022 г.).

Статистический анализ

Принципы расчёта размера выборки: размер выборки предварительно не рассчитывался.

Методы статистического анализа данных: статистический анализ проводился с помощью программ STATISTICA 12.0 («StatSoft», США) и Epi Info версия 7.2.1.0 (CDC, США). Сравнение между категориальными переменными, выраженными в процентах (%), выполнялись с помощью критерия согласия Пирсона «хи-квадрат» (χ^2) или точного критерия Фишера, в зависимости от обстоятельств [15]. Для оценки влияния признака в узлах моделей использовалось определение отношения шансов (ОШ) и 95% доверительного интервала (95% ДИ). Статистическая значимость была установлена на уровне $p < 0,05$.

Работа была выполнена на языке программирования Python версии 3.7.13 с использованием встроенных библиотек (модулей) для работы с массивами данных, машинного обучения и визуализации [12, 16]. Для настройки алгоритма модели дерева решений был проведён подбор оптимальных гиперпараметров: в качестве критерия разбиения и меры неопределённости для набора данных D в узле дерева решений была выбрана энтропия $Info(D)$, максимальная глубина дерева, то есть количество узлов в дереве, составила 15, а пороговое значения критерия разбиения для разделения данных в узле было равно 0,1.

Энтропия $Info(D)$ рассчитывается по формуле:

$$Info(D) = -\sum_{i=1}^n p(i|t) \times \log_2 p(i|t),$$

где D — это набор данных, который содержит количество n различных классов C_i ($i=1, 2, 3, \dots, n$); $p(i)$ — вероятность (probability) или доля образцов, принадлежащая классу C_i для узла t [16, 17]. В модели для больных с эритемой $n=2$ (классы: больные СИ ЭФ ИКБ с КЭ и больные МИ ЭФ ИКБ), а в дереве решений для пациентов, у которых эритема отсутствует, $n=3$ (классы: больные СИ БФ ИКБ с КЭ, больные МИ БФ ИКБ и больные МИ ЛФ КЭ).

В использованном нами алгоритме ID3 дерево решений строится рекурсивно путём выбора предикторов в соответствии с порядком (правилом), основанным на критерии получения и расщепления данных в каждом узле (node), который определяется мерой прироста информации (information gain). Прирост информации зависит от степени неопределённости или разнородности (impurity) классов в узле и представляет ожидаемое снижение энтропии $InfoA(D)$ при каждом расщеплении данных

по признаку. Чем меньше энтропия, тем меньше разнородности или случайности в системе. На каждом узле древовидный алгоритм выбирает точку разбиения (split value) с наибольшим приростом информации, которая минимизирует неопределённость исхода в каждом узле и делит данные предикторной переменной на те, где она меньше или больше определённого порогового значения. Когда никакие дальнейшие разбиения сделать невозможно, дерево считается полностью выращенным, а каждый терминальный узел или лист (leaf) имеет записи с единственным классом, то есть вариантом диагноза в данном исследовании. Входные переменные (input), то есть исходные данные, включённые в модели, были категориальными, имеющими бинарный отклик по типу «Да»=1 или «Нет»=0, или количественными (числовыми) [16, 17].

Древовидные модели позволяют оценить важность предикторов. Важность признака (feature importance, FI) может дать представление о том, какие переменные имеют наиболее сильный эффект в моделях классификации. Для определения важности признаков был использован модуль Sklearn пакета библиотек Python, имеющий в модели «Леса деревьев» («Forests of trees») атрибут «Feature importances», который показывает, какое разбиение переменных наиболее эффективно позволяет различать изученные классы [18].

Оценка точности и эффективности моделей была основана на определении площади под ROC-кривой (AUC, от англ. area under receiver operating characteristic curve) и на показателях, отражающих чувствительность или полноту (sensitivity или recall), специфичность

(specificity), точность (accuracy), прецизионность или прогностическую ценность положительного результата (precision или positive predictive value), а также на показателе F1 (F1 score), которые были рассчитаны из матрицы путаницы, то есть таблицы прогнозируемых и фактических значений классификатора [18, 19]. Показатели AUC были представлены как $M \pm SD$, где M — среднее значение, а SD — стандартное отклонение.

РЕЗУЛЬТАТЫ

Объекты (участники) исследования

Среди 211 больных клещевыми инфекциями без эритемы был проведён поиск клинико-лабораторных предикторов, позволяющих различать между собой три класса пациентов в зависимости от окончательного диагноза: СИ БФ ИКБ с ЛФ КЭ (класс 1, $n=38$, средний возраст: $48,53 \pm 3,41$ лет), МИ БФ ИКБ (класс 2, $n=93$, средний возраст: $47,10 \pm 3,12$ лет) или МИ ЛФ КЭ (класс 3, $n=80$, средний возраст: $45,33 \pm 3,42$ лет). Среди 80 пациентов с эритемой анализ предикторных переменных позволил проводить дифференциальный диагноз между двумя классами больных: СИ ЭФ ИКБ с ЛФ КЭ (класс 4, $n=37$, средний возраст: $47,36 \pm 4,1$ лет) и МИ ЭФ ИКБ (класс 5, $n=43$, средний возраст: $49,21 \pm 3,20$ лет).

Каждый из классов пациентов был разделён на обучающую и тестовую выборки в отношении 70 и 30% (табл. 1). В каждом из классов состав обучающей и тестовой выборок был сопоставим в зависимости от пола и возраста больных.

Таблица 1. Численность и распределение в зависимости от пола в обучающей и тестовой выборках больных, имевших диагнозы различных клинических форм клещевых инфекций с наличием или отсутствием мигрирующей эритемы на месте присасывания иксодового клеща

Table 1. Number and sex-depending distribution in the training and test datasets of patients diagnosed with different clinical forms of tick-borne infections with or without erythema migrans at the site of ixodid tick bite

Выборка	Пол	Численность групп, абс. (%)				
		Больные клещевыми инфекциями без эритемы			Больные клещевыми инфекциями с эритемой	
		Класс 1 СИ $n=38$	Класс 2 МИ ИКБ $n=93$	Класс 3 МИ КЭ $n=80$	Класс 4 СИ $n=37$	Класс 5 МИ ИКБ $n=43$
Обучающая	Всего	27 (71,0)	65 (70,0)	56 (70,0)	26 (70,3)	30 (69,8)
	мужчины	15 (55,6)	31 (47,7)	30 (53,6)	16 (61,5)	15 (50,0)
	женщины	12 (44,4)	34 (52,3)	26 (46,4)	10 (38,5)	15 (50,0)
Тестовая	Всего	11 (29,0)	28 (30,0)	24 (30,0)	11 (29,7)	13 (30,2)
	мужчины	6 (54,5)	13 (46,4)	13 (54,2)	6 (54,5)	7 (53,8)
	женщины	5 (45,5)	15 (53,6)	11 (45,8)	5 (45,5)	6 (46,2)

Примечание. СИ — смешанная инфекция; МИ — моноинфекция; ИКБ — иксодовый клещевой боррелиоз; КЭ — клещевой энцефалит.

Note: СИ (Mixl) — mixed infection; МИ (Monl) — mono-infection; ИКБ (LB) — Lyme borreliosis; КЭ (TBE) — tick-borne encephalitis.

Основные результаты исследования

При построении моделей деревьев решений для ранней дифференциальной диагностики у пациентов с подозрением на изолированные и смешанные формы клещевых инфекций, имевших или не имевших мигрирующую эритему на месте присасывания иксодового клеща, в анализ было включено 55 входных предикторных переменных, полученных при клиническом и лабораторном обследовании больных в первую неделю заболевания до верификации окончательного диагноза. В результате были установлены нелинейные связи между значениями определённых предикторов и классами, в качестве которых служили варианты окончательных диагнозов пациентов. В исследовании для построения деревьев решений использовались предикторные переменные, которые имели наиболее высокие значения важности для дифференциальной диагностики смешанной патологии и моноинфекций в группах больных клещевыми инфекциями.

С помощью определения показателей относительной важности каждого признака было установлено, что для ранней дифференциальной диагностики между тремя классами пациентов, не имевших мигрирующей эритемы, с диагнозами СИ БФ ИКБ с КЭ, МИ БФ ИКБ или МИ ЛФ КЭ, наиболее сильной дискриминационной способностью обладали следующие предикторы: максимальная высота лихорадки ($FI=0,26$), СОЭ ($FI=0,16$), ИСНМ ($FI=0,15$), озноб ($FI=0,14$), абсолютное количество Ig ($FI=0,14$), абсолютное количество RE-LYMP ($FI=0,11$) и относительное число эозинофилов (ЕО) ($FI=0,05$) в гемограмме.

Наиболее важными переменными для дифференциальной диагностики между больными эритемными формами изолированно протекающего ИКБ и СИ ИКБ с КЭ были следующие четыре атрибута в порядке убывания значимости: максимальная высота лихорадки ($FI=0,34$), абсолютное количество Ig ($FI=0,28$), относительное число базофилов (BASO) ($FI=0,21$) и абсолютное количество RE-LYMP ($FI=0,19$) в гемограмме.

Модель дерева решений для ранней дифференциальной диагностики у пациентов с подозрением на клещевые инфекции, не имеющих мигрирующей эритемы на месте присасывания иксодового клеща, представлена на рис. 1.

Корневой узел обучающей выборки этой модели включал три класса больных с СИ БФ ИКБ и КЭ, с МИ БФ ИКБ и с МИ ЛФ КЭ, поэтому энтропия в первом узле была равна 1,50. В качестве начальной предикторной переменной в корневом узле была выбрана максимальная высота лихорадки с пороговым значением меньше / равно или выше $38,0^\circ\text{C}$. Таким образом, эта модель дерева решений имела две основные ветви для классификации пациентов, не имеющих мигрирующей эритемы. В ветви дерева с максимальной температурой тела у пациентов

ниже или равно порогового значения $38,0^\circ\text{C}$ в точке расщепления последующие четыре узла модели включали предикторы: ИСНМ с пороговым значением меньше / равно или больше 9,50 усл. ед., СОЭ меньше / равно или больше 9,50 мм/ч, RE-LYMP меньше / равно или больше $0,075 \times 10^9/\text{л}$, а также относительное число ЕО меньше / равно или больше 1,60%.

В каждом из узлов определена статистическая значимость различий и рассчитаны показатели ОШ для преобладающего по численности класса по сравнению с остальными пациентами, имеющими два других варианта диагноза клещевых инфекций.

Из 95 пациентов, имеющих субфебрильную лихорадку или нормальную температуру тела, были представлены все три класса, но преобладали пациенты с диагнозом МИ БФ ИКБ — 61 больной, или 64,2% ($OШ=21,98$ (7,30–66,17), $p < 0,001$). При условии выполнения двух последующих правил: ИСНМ $\leq 9,50$ усл. ед. и СОЭ $\leq 9,50$ мм/ч, этот диагноз в конечном узле был у 43 больных, или у 86,0% ($OШ=7,29$ (2,58–20,62), $\chi^2=15,74$, $p < 0,001$), а энтропия снижалась до 0,58. Если у пациентов, имеющих субфебрильную температуру тела, соблюдалось правило: ИСНМ $\leq 9,50$ усл. ед., СОЭ $> 9,50$ мм/ч и ЕО $\leq 1,60\%$, то в конечном узле 100% больных имели диагноз ЛФ КЭ, а энтропия снижалась до 0. Если у больных относительное число ЕО в периферической крови было больше 1,60%, то наиболее вероятным диагнозом была МИ БФ ИКБ, который имели 16 (88,89%) пациентов, а энтропия была равна 0,50. Среди пациентов с максимальной температурой тела, не превышающей $38,0^\circ\text{C}$, которые имели показатель ИСНМ $> 9,50$ усл. ед., преобладали больные с СИ БФ ИКБ и КЭ, составляя в узле 8 (80,0%) пациентов ($OШ=52,0$ (8,94–301,64), $p < 0,001$). Сочетание субфебрильной лихорадки, ИСНМ $> 9,50$ усл. ед. и RE-LYMP $\leq 0,075 \times 10^9/\text{л}$ было характерно исключительно для больных с диагнозом СИ БФ ИКБ и КЭ, то есть у 8 (100%) пациентов в конечном узле. Если пациенты имели показатель RE-LYMP $> 0,075 \times 10^9/\text{л}$, то в 100% это соответствовало классу больных МИ БФ ИКБ. Поэтому показатели энтропии в этих двух конечных узлах были равны 0.

Ветвь дерева с максимальной высотой лихорадки у пациентов выше порогового значения $38,0^\circ\text{C}$ в точке расщепления включала еще две последующие предикторные переменные: наличие или отсутствие у пациента симптома озноба, а также абсолютное число Ig в гемограмме с пороговым значением меньше / равно или больше $0,025 \times 10^9/\text{л}$. В целом из 53 пациентов с фебрильной лихорадкой 36 (67,92%) больных имели диагноз МИ ЛФ КЭ ($OШ=7,94$ (3,71–16,96), $\chi^2=31,78$, $p < 0,001$). Если у пациентов фебрильная лихорадка сопровождалась ознобом, то этот диагноз был уже у 25, или 96,15%, то есть у подавляющего большинства больных ($OШ=36,36$ (4,27–309,43), $p < 0,001$), а энтропия в конечном узле уменьшалась до 0,24. В случае если у пациентов с фебрильной температурой озноб

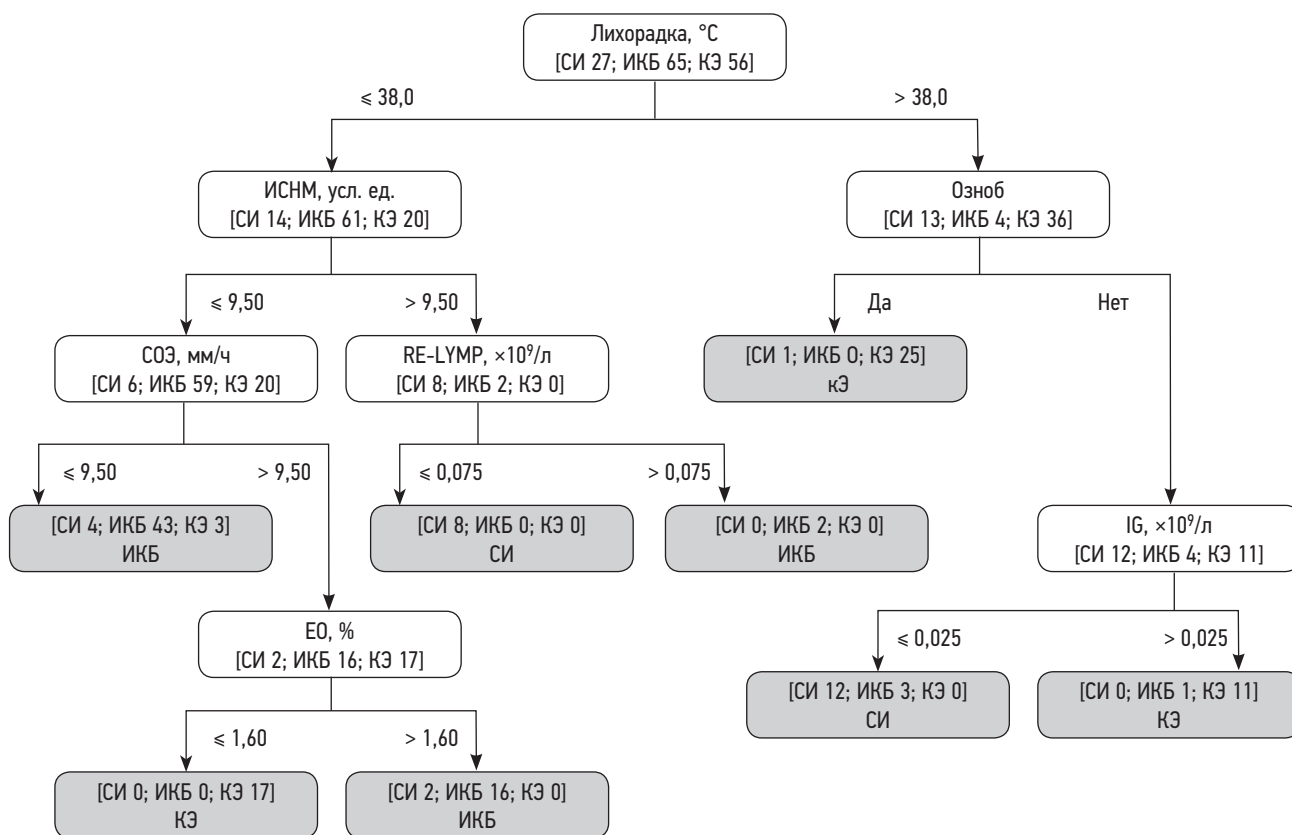


Рис. 1. Алгоритм дерева решения для ранней клинической дифференциальной диагностики между тремя классами пациентов без мигрирующей эритемы на месте присасывания иксодового клеща, имевших диагнозы смешанной инфекции безэритемной формы иксодового клещевого боррелиоза и клещевого энцефалита, моноинфекции безэритемной формы иксодового клещевого боррелиоза или моноинфекции клещевого энцефалита: ИКБ — моноинфекция иксодового клещевого боррелиоза; КЭ — моноинфекция клещевого энцефалита; СИ — смешанная инфекция иксодового клещевого боррелиоза и клещевого энцефалита; ИСНМ — индекс соотношения нейтрофилов и моноцитов, усл. ед.; СОЭ — скорость оседания эритроцитов, мм/ч; ЕО — относительное число эозинофилов, %; RE-LYMP — абсолютное число реактивных лимфоцитов, $\times 10^9/\text{л}$; IG — абсолютное число незрелых гранулоцитов, $\times 10^9/\text{л}$.

Fig. 1. Decision tree model for early clinical differential diagnosis between the following three classes of patients without erythema migrans at the site of ixodid tick bite: the mixed infection of Lyme borreliosis non-erythematous form and tick-borne encephalitis, the mono-infection of Lyme borreliosis non-erythematous form or the mono-infection of tick-borne encephalitis: ИКБ — mono-infection of Lyme borreliosis; КЭ — mono-infection of tick-borne encephalitis; СИ — mixed infection of Lyme borreliosis and tick-borne encephalitis; ИСНМ — neutrophil-monocyte ratio, units; СОЭ — erythrocyte sedimentation rate, mm/h; ЕО — number of eosinophils, %; RE-LYMP — absolute number of reactive lymphocytes, $\times 10^9/\text{l}$; IG — absolute number of immature granulocytes, $\times 10^9/\text{l}$.

отсутствовал, а параметр $IG \leq 0,025 \times 10^9/\text{л}$, то доминирующим классом были больные с СИ БФ ИКБ и КЭ, составляя 12, или 80,0%, пациентов. Энтропия в этом узле равнялась 0,72. Если у больных показатель $IG > 0,025 \times 10^9/\text{л}$, то в этом конечном узле преобладали пациенты с МИ ЛФ КЭ, то есть 11, или 91,67%, больных. Энтропия в конечном узле составила 0,40. Таким образом, эта модель дерева решений имела восемь конечных узлов, в трёх из которых преобладающим классом были больные МИ БФ ИКБ, в трёх других преобладал диагноз МИ ЛФ КЭ, а в оставшихся двух узлах были повышена вероятность СИ БФ ИКБ с ЛФ КЭ.

На рис. 2 представлена модель дерева решений для ранней клинической дифференциальной диагностики между пациентами с МИ ЭФ ИКБ и СИ ЭФ ИКБ с КЭ, полученная на обучающей выборке.

Как и в случае больных клещевыми инфекциями, не имеющих эритемы, в качестве предикторной переменной в корневом узле этой модели была выбрана максимальная высота лихорадки. Энтропия в первом узле была равна 1,0. Если значения этого предиктора были ниже или равно порогового значения $37,40^\circ\text{C}$ в точке расщепления, то у больных, имеющих эритему, в 15,13 (3,97–57,64) раза преобладали шансы диагноза МИ ЭФ ИКБ (у 22 больных, или у 84,61%, в узле) над сочетанным течением ЭФ ИКБ и КЭ ($p < 0,001$). В ветви дерева с максимальной температурой тела у пациентов, имеющих эритему ниже или равно порогового значения $37,40^\circ\text{C}$ в точке расщепления, последующим предиктором был показатель IG с пороговым значением меньше / равно или больше $0,005 \times 10^9/\text{л}$. При условии соблюдения правила: температура тела ниже или равно $37,40^\circ\text{C}$ и параметр

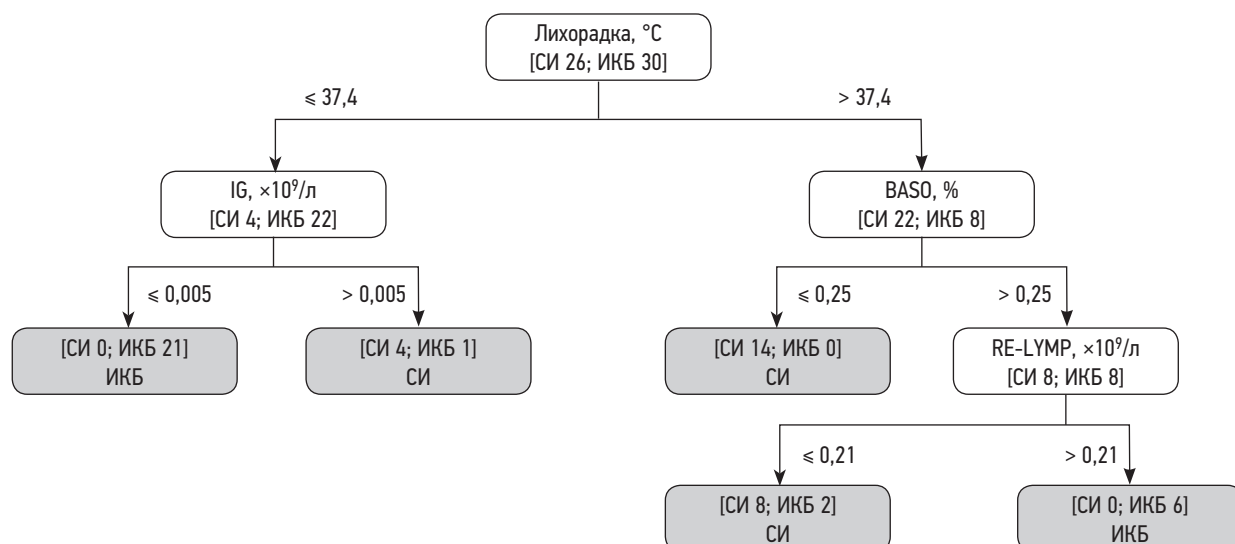


Рис. 2. Алгоритм дерева решения для ранней клинической дифференциальной диагностики между двумя классами пациентов с мигрирующей эритемой на месте присасывания иксодового клеща, имевших диагнозы смешанной инфекции эритемной формы иксодового клещевого боррелиоза и клещевого энцефалита или моноинфекции эритемной формы иксодового клещевого боррелиоза: BASO — относительное число базофилов, %; IG — абсолютное число незрелых гранулоцитов, $\times 10^9/\text{л}$; RE-LYMP — абсолютное число реактивных лимфоцитов, $\times 10^9/\text{л}$.

Fig. 2. Decision tree model for early clinical differential diagnosis between the following two classes of patients with erythema migrans at the site of ixodid tick bite: the mixed infection of Lyme borreliosis erythematous form and tick-borne encephalitis or the mono-infection of Lyme borreliosis erythematous form: BASO — number of basophils, %; IG — absolute number of immature granulocytes, $\times 10^9/\text{l}$; RE-LYMP — absolute number of reactive lymphocytes, $\times 10^9/\text{l}$.

$IG \leq 0,005 \times 10^9/\text{л}$, 100% больных имели диагноз МИ ЭФ ИКБ, а энтропия была равна 0. В случае если у пациентов показатель $IG > 0,005 \times 10^9/\text{л}$, то наиболее вероятным диагнозом у 80,0% больных была СИ ЭФ ИКБ и ЛФ КЭ, а энтропия составила 0,72.

Среди пациентов с лихорадкой выше 37,4 °C доминировал диагноз СИ ЭФ ИКБ с КЭ, то есть у 22 больных, или 73,33%, в узле, что свидетельствовало об относительно более выраженном лихорадочном синдроме у этой группы. В этой ветви алгоритма дискриминационными переменными во втором и третьем узлах были относительное число BASO и абсолютное число RE-LYMP в периферической крови. В случае соблюдения правила: $BASO \leq 0,25\%$ единственным классом в конечном узле были 14 пациентов с диагнозом СИ ЭФ ИКБ с ЛФ КЭ, а энтропия была равна 0. Если у пациентов с эритемой значение этого показателя в точке расщепления было больше 0,25%, то больные с равной вероятностью в 50% могли иметь диагнозы микст- или моноинфекции, что дополнительно потребовало включения в модель еще одной переменной в третьем узле дерева решений — абсолютного числа RE-LYMP. В случае соблюдения правила: $BASO > 0,25\%$ и $RE-LYMP \leq 0,21 \times 10^9/\text{л}$ в крови, наиболее высокую вероятность имел класс пациентов с диагнозом СИ ЭФ ИКБ и ЛФ КЭ, то есть у 8, или 80,0%, больных. В случае если уровень RE-LYMP превышал $0,21 \times 10^9/\text{л}$, то единственным классом в конечном узле были больные с МИ ЭФ ИКБ.

Энтропия в первом случае не превышала 0,72, а во втором была равна 0. В результате алгоритм дерева решений заканчивался пятью узлами, в трёх из которых преобладающим классом были больные СИ ЭФ ИКБ и ЛФ КЭ, в двух других была повышена вероятность МИ ЭФ ИКБ.

Дополнительные результаты исследования

Прогностическая ценность предикторов и правил обоих алгоритмов деревьев решений для дифференциального диагноза клещевых инфекций была подтверждена на тестовой группе. В табл. 2 приведена матрица путаницы тестовой выборки, использованная для оценки точности моделей деревьев решений для ранней дифференциальной диагностики больных клещевыми инфекциями с наличием или отсутствием мигрирующей эритемы на месте присасывания иксодового клеща. Результаты оценки эффективности построенных моделей деревьев решений после валидации с применением тестовой выборки представлены в табл. 3. Обе модели деревьев решений получили высокую прогностическую оценку на основании определения различных метрик, таких как чувствительность, специфичность, точность, прецизионность и показатель F1. В частности, с применением тестовой группы была подтверждена чувствительность и специфичность модели 1 дерева решений для прогноза для трех классов больных без эритемы: СИ БФ ИКБ с КЭ (72,72 и 92,31%), МИ БФ ИКБ (89,29 и 91,43%) и МИ ЛФ КЭ (83,33 и 92,31%), а также модели 2 для дифференциального

Таблица 2. Матрица путаницы, показывающая точность моделей деревьев решений для ранней дифференциальной диагностики больных с различными клиническими формами клещевых инфекций с наличием или отсутствием мигрирующей эритемы на месте присасывания иксодового клеща

Table 2. Confusion matrix depicting the accuracy of the decision tree models for early differential diagnosis of patients with different clinical forms of tick-borne infections with or without erythema migrans at the site of ixodid tick bite

Ожидаемые результаты	Фактические результаты				
	Модель 1 для больных клещевыми инфекциями без эритемы			Модель 2 для больных клещевыми инфекциями с эритемой	
	Класс 1 СИ	Класс 2 МИ ИКБ	Класс 3 МИ КЭ	Класс 4 СИ	Класс 5 МИ ИКБ
Класс 1 СИ	8	2	2	9	2
Класс 2 МИ ИКБ	1	25	2	2	11
Класс 3 МИ КЭ	2	1	20	–	–
Всего	11	28	24	11	13

Примечание. СИ — смешанная инфекция; МИ — моноинфекция; ИКБ — иксодовый клещевой боррелиоз; КЭ — клещевой энцефалит.

Note: СИ (Mixl) — mixed infection; МИ (Monl) — mono-infection; ИКБ (LB) — Lyme borreliosis; КЭ (TBE) — tick-borne encephalitis.

Таблица 3. Оценка точности моделей деревьев решений для ранней дифференциальной диагностики больных с различными клиническими формами клещевых инфекций с наличием или отсутствием мигрирующей эритемы на месте присасывания иксодового клеща

Table 3. Accuracy evaluation of decision tree models for early differential diagnosis of patients with different clinical forms of tick-borne infections with or without erythema migrans at the site of ixodid tick bite

Критерий оценки	Модель 1 для больных клещевыми инфекциями без эритемы			Модель 2 для больных клещевыми инфекциями с эритемой	
	Класс 1 СИ n=11	Класс 2 МИ ИКБ n=28	Класс 3 МИ КЭ n=24	Класс 4 СИ n=11	Класс 5 МИ ИКБ n=13
Чувствительность, %	72,72	89,29	83,33	81,81	84,62
Специфичность, %	92,31	91,43	92,31	84,61	81,81
Точность, %	88,89	90,48	88,89	83,33	83,33
Прецизионность, %	66,67	89,29	86,96	81,81	84,62
Показатель F1, %	69,56	89,29	85,11	81,81	84,62
AUC, M±SD	0,91±0,15	0,93±0,12	0,95±0,13	0,94±0,06	0,94±0,06

Примечание. СИ — смешанная инфекция; МИ — моноинфекция; ИКБ — иксодовый клещевой боррелиоз; КЭ — клещевой энцефалит.

Note: СИ (Mixl) — mixed infection; МИ (Monl) — mono-infection; ИКБ (LB) — Lyme borreliosis; КЭ (TBE) — tick-borne encephalitis.

диагноза между двумя классами пациентов с эритемой: СИ ЭФ ИКБ с КЭ (81,81 и 84,61%) и МИ ЭФ ИКБ (84,62 и 81,81%), что свидетельствовало об отсутствии переобучения и применимости полученных алгоритмов деревьев решений для других сопоставимых выборок.

Для оценки качества работы обученных моделей по предикторам из тестовых выборок был спрогнозирован диагноз каждого из пациентов. На основе реальных

и предсказанных значений были построены ROC-кривые с вычислением AUC, которые показывают качество моделей. В результате оба алгоритма деревьев решений продемонстрировали отличные результаты дискриминационной способности правильно распределять пациентов по соответствующим классам с различными диагнозами, поскольку AUC во всех случаях был больше 0,90 (см. табл. 3).

ОБСУЖДЕНИЕ

Резюме основного результата исследования

С использованием клиничко-лабораторных параметров разработаны два алгоритма дерева решений, имеющие высокую чувствительность, специфичность и точность для ранней клинической дифференциальной диагностики между изолированными и смешанными формами острого ИКБ и КЭ с преобладанием в клинической картине лихорадочно-интоксикационного синдрома. Алгоритм дерева решений для дифференциальной диагностики в начале заболевания между больными клещевыми инфекциями без эритемы, которые имели диагнозы СИ БФ ИКБ с КЭ, МИ БФ ИКБ или МИ ЛФ КЭ, включал следующие наиболее важные предикторы: максимальную высоту лихорадки, озноб, ИСНМ, СОЭ, абсолютное число RE-LYMP и IG, процентное содержание эозинофилов. Модель для дифференциального диагноза между пациентами с эритемой, которые имели диагнозы СИ ЭФ ИКБ с КЭ или МИ ЭФ ИКБ, включала: максимальную высоту лихорадки, абсолютное число RE-LYMP и IG, а также процентное содержание базофилов.

Обсуждение основного результата исследования

Клиническая картина СИ ИКБ с КЭ отличается крайним многообразием, что во многом зависит от применения различных классификаций, а также от формы и стадии течения этих заболеваний, поэтому дифференциальный диагноз клещевых инфекций, в особенности в первую неделю заболевания, может быть сложен и требует принципиально новых, ранее не применявшихся подходов [3–5]. В литературе имеются единичные исследования, посвящённые построению моделей множественной линейной или логистической регрессии для прогнозирования исходов ИКБ (Лайм-боррелиозов) на основе анализа клинических или иммунологических данных [20, 21].

Алгоритмы деревьев решений в настоящее время применяются для клинической дифференциальной диагностики и стратегий медицинской сортировки больных с различными инфекционными заболеваниями, сопровождающимися синдромом лихорадки. В частности, клиничко-лабораторные данные, включая параметры гемограммы, используются в моделях деревьев решений для ранней дифференциальной диагностики лихорадки денге и коронавирусной инфекции COVID-19 от других заболеваний [8, 10]. Кроме того, разработаны модели деревьев решений для прогноза тяжёлого течения и/или летального исхода лихорадки денге и COVID-19, что в условиях эпидемий имеет особое значение не только для поддержки принятия клинических решений, но и для планирования распределения ресурсов здравоохранения [9, 11].

Показано, что хотя микст-инфекция ИКБ с КЭ в начальный период заболевания не имеет специфических симптомов, свойственных сочетанному заболеванию, однако, как правило, протекает клинически более тяжело за счёт выраженных проявлений синдромов лихорадки и интоксикации по сравнению с моноинфекцией острого ИКБ, но более легко по сравнению с изолированной формой КЭ, что нашло подтверждение в построенных нами алгоритмах деревьев решений, поскольку такие параметры, как максимальная высота лихорадки и/или озноб, являлись наиболее важными предикторами [4–6]. Кроме того, известно, что большинство исследователей считают, что БФ острого ИКБ имеет более тяжёлое течение, чем ЭФ [22, 23].

Среди лабораторных параметров, отражающих интенсивность общевоспалительного синдрома, имеющих дифференциальное значение в разработанных нами моделях деревьев решений, были показатели ИСНМ, СОЭ, абсолютного числа IG и/или RE-LYMP. Известно, что большинство больных как СИ ИКБ с КЭ, так и МИ ЛФ КЭ в начале заболевания имели тенденцию к умеренному повышению уровня СОЭ и количества нейтрофилов, включая число палочкоядерных и незрелых форм гранулоцитов в лейкоцитарной формуле периферической крови [2, 4, 23], что согласуется с полученными нами данными. Кроме того, снижение показателя RE-LYMP у больных СИ ИКБ с КЭ по сравнению с МИ ИКБ, возможно, отражает уменьшение числа активированных Т-лимфоцитов в иммунном статусе в первые дни болезни [3, 24]. Повышение числа незрелых гранулоцитов IG и лейкоцитарного индекса ИСНМ, а также снижение параметра RE-LYMP рассматривают как важные маркеры тяжести и/или повышенного риска летального исхода у больных с COVID-19 [25–27] или с лихорадкой денге [28].

Изменения двух других параметров гемограммы в построенных нами моделях — числа эозинофилов или базофилов, по-видимому, связаны с различиями в иммунном ответе на вирус и боррелии. Хотя известно, что эти показатели при клещевых инфекциях в большинстве случаев остаются в пределах нормы, но при МИ ИКБ, в особенности в случае ЭФ, эти параметры имеют существенную тенденцию к росту в первые дни болезни по сравнению со здоровыми людьми и больными с микст- и моноинфекцией КЭ [22, 23], что было подтверждено в настоящем исследовании.

Ограничения исследования

Ограничениями данного исследования является то, что модели деревьев решений были получены на относительно небольших выборках и обучены на данных только одной когорты. Кроме того, данные модели могут быть применимы только для взрослых пациентов, а также для больных с изолированной или смешанной формой острого ИКБ и/или ЛФ КЭ, которые были госпитализированы в первую

неделю болезни и не имели признаков поражения нервной системы.

ЗАКЛЮЧЕНИЕ

Наиболее распространённые в России природно-очаговые инфекции с трансмиссивным механизмом передачи — КЗ и ИКБ — нередко протекают в виде смешанной формы, которую зачастую клинически сложно отличить от моноинфекции в начале болезни, что также может быть обусловлено поздней лабораторной верификацией диагноза и требует поиска принципиально нового подхода к проблеме раннего дифференциального диагноза клещевых инфекций. Одним из таких подходов являются модели деревьев решений, к основным преимуществам которых по сравнению с методами регрессионного анализа можно отнести наглядность и простоту интерпретации при практическом применении. Алгоритмы деревьев решений, построенные с использованием данных клинико-лабораторного обследования пациентов, в настоящее время применяются для ранней дифференциальной диагностики, прогнозирования тяжёлого течения и/или летального исхода различных инфекционных заболеваний, включая лихорадку денге и COVID-19.

В результате настоящего исследования с помощью метода машинного обучения с применением языка программирования Python впервые были построены две модели деревьев решений для ранней дифференциальной диагностики больных с изолированной или смешанной формой острого ИКБ и КЗ с преобладанием в клинической картине синдрома лихорадки на основе данных анализа наиболее важных клинико-лабораторных предикторов в первую неделю болезни. Один алгоритм дерева решений предназначен для клинического дифференциального диагноза между больными, не имеющими мигрирующей эритемы на месте присасывания иксодового клеща, — СИ БФ ИКБ с КЗ, МИ БФ ИКБ или МИ ЛФ КЗ, а другой — для дифференциальной диагностики между пациентами с эритемой, имеющими СИ ЭФ ИКБ с КЗ или МИ ЭФ ИКБ. Обе модели деревьев решений продемонстрировали высокие показатели оценки точности, включавшие чувствительность, специфичность, точность, прецизионность и показатель F1, а также площадь AUC под ROC-кривой, которая превышала 0,90.

Данное исследование позволяет полагать, что алгоритмы классификации, использующие простые клинические и гематологические параметры, могут иметь практическое применение для определения тактики ведения пациентов

и обоснования показаний для выбора адекватной этиотропной терапии клещевых инфекций в первые дни болезни до получения результатов лабораторной верификации диагноза.

ДОПОЛНИТЕЛЬНО

Источник финансирования. Исследование выполнено за счёт гранта Российского научного фонда № 22-15-20010, <https://rscf.ru/project/22-15-20010/> и средств Администрации Томской области.

Конфликт интересов. Авторы данной статьи подтвердили отсутствие конфликта интересов, связанного с публикацией настоящей статьи.

Вклад авторов. Все авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции, проведение исследования и подготовку статьи, прочли и одобрили финальную версию перед публикацией). Наибольший вклад распределён следующим образом: Е.Н. Ильинских — формирование концепции и дизайна исследования, правка рукописи статьи, формулировка выводов и интерпретация результатов исследования; Е.Н. Филатова — подбор и одобрение клинического материала исследования в соответствии с этическими нормами, обработка архивных историй болезни, статистическая обработка полученных данных, написание статьи; К.В. Самойлов — подбор и анализ литературных источников, написание статьи; А.В. Семёнова — обработка архивных историй болезни; С.В. Аксёнов — разработка математических моделей, статистическая обработка материалов, написание статьи.

ADDITIONAL INFORMATION

Funding source. The study was supported by the grant of the Russian Science Foundation No. 22-15-20010, <https://rscf.ru/project/22-15-20010/> and the Tomsk Region Administration.

Competing interests. The authors declare that they have no competing interests.

Authors' contribution. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work. E.N. Ilyinskikh — formation of the concept and design of the study, editing manuscript, formulation of conclusions and interpretation of the results of the study; E.N. Filatova — selection and approval of clinical research material in accordance with ethical standards, processing of archival patient records, statistical processing of materials, writing the manuscript; K.V. Samoylov — collecting and analyzing literature data, writing the manuscript; A.V. Semenova — processing of archival patient records; S.V. Axyonov — development of mathematical models, statistical processing of materials, writing the manuscript.

СПИСОК ЛИТЕРАТУРЫ

1. Лобзин Ю.В., Усков А.Н., Козлов С.С. Лайм-боррелиоз: иксодовые клещевые боррелиозы. Санкт-Петербург : Фолиант, 2000. 156 с.

2. Иерусалимский А.П. Клещевой энцефалит : руководство для врачей. Новосибирск : Государственная медицинская академия МЗ РФ, 2001. 360 с.

3. Бондаренко А.Л., Зыкова И.В., Аббасова С.В., Тихомолова Е.Г., Нехорошкина Е.Л. Микст-инфекция клещевого энцефалита и иксодовых клещевых боррелиозов // Инфекционные болезни. 2011. Т. 9, № 4. С. 54–63.
4. Субботин А.В., Семенов В.А., Этенко Д.А. Проблема современных смешанных нейроинфекций, передающихся иксодовыми клещами // Архивъ внутренней медицины. 2012. № 2 (4). С. 35–39.
5. Конькова-Рейдман А.Б., Злобин В.И. Клинический полиморфизм иксодовых клещевых боррелиозов (микст-инфекция с клещевым энцефалитом) на территории Южно-Уральского региона России // Сибирский медицинский журнал (Иркутск). 2011. Т. 100, № 1. С. 17–20.
6. Миноранская Н.С., Миноранская Е.И. Клинико-эпидемиологическая характеристика микст-инфекции клещевого боррелиоза и клещевого энцефалита в Красноярском крае // Казанский медицинский журнал. 2013. Т. 94, № 2. С. 211–215.
7. Андропова Н.В., Миноранская Н.С., Миноранская Е.И. Специфический иммунный ответ и некоторые отдаленные результаты при остром течении иксодового клещевого боррелиоза и микст-инфекции клещевого энцефалита и иксодового клещевого боррелиоза // Сибирский медицинский журнал (Иркутск). 2011. Т. 100, № 1. С. 54–57.
8. Tanner L., Schreiber M., Low J.G., et al. Decision tree algorithms predict the diagnosis and outcome of dengue fever in the early phase of illness // PLoS Negl Trop Dis. 2008. Vol. 2, N 3. P. e196. doi: 10.1371/journal.pntd.0000196
9. Tamibmaniam J., Hussin N., Cheah W.K., Ng K.S., Muninathan P. Proposal of a clinical decision tree algorithm using factors associated with severe dengue infection // PLoS One. 2016. Vol. 11, N 8. P. e0161696. doi: 10.1371/journal.pone.0161696
10. Dobrijević D., Andrijević L., Antić J., Rakić G., Pastor K. Hemogram-based decision tree models for discriminating COVID-19 from RSV in infants // J Clin Lab Anal. 2023. Vol. 37, N 6. P. e24862. doi: 10.1002/jcla.24862
11. Kumaran M., Pham T.M., Wang K., et al. Predicting the risk factors associated with severe outcomes among COVID-19 patients—decision tree modeling approach // Front Public Health. 2022. Vol. 10. P. 838514. doi: 10.3389/fpubh.2022.838514
12. Брюс П., Брюс Э. Практическая статистика для специалистов Data Science. Санкт-Петербург : БХВ-Петербург, 2018. 304 с.
13. Kowalska-Kępczyńska A., Mleczko M., Domerecka W., Krasowska D., Donica H. Assessment of immune cell activation in pemphigus // Cells. 2022. Vol. 11, N 12. P. 1912. doi: 10.3390/cells11121912
14. Гребенникова И.В., Лидохова О.В., Макеева А.В., и др. Возрастные аспекты изменения лейкоцитарных индексов при COVID-19 // Научно-медицинский вестник Центрального Черноземья. 2022. № 87. С. 9–15.
15. Петри А., Сэбин К. Наглядная медицинская статистика / под ред. В.П. Леонова. Москва : ГЭОТАР-Медиа, 2015. 216 с.
16. Huang Y., Li S., Lin B., et al. Early Detection of college students' psychological problems based on decision tree model // Front Psychol. 2022. Vol. 13. P. 946998. doi: 10.3389/fpsyg.2022.946998
17. Mukhopadhyay S. Advanced data analytics using Python: with machine learning, deep learning and NLP examples. Kolkata : Apress Berkeley, 2018. 186 p. doi: 10.1007/978-1-4842-3450-1
18. Kalafi E.Y., Nor N.A.M., Taib N.A., et al. Machine learning and deep learning approaches in breast cancer survival prediction using clinical data // Folia Biol (Praha). 2019. Vol. 65, N 5–6. P. 212–220.
19. Лакшманан В., Робинсон С., Мунн М. Машинное обучение. Паттерны проектирования. Санкт-Петербург : БХВ-Петербург, 2022. 448 с.
20. Turk S.P., Lumbard K., Liepshutz K., et al. Post-treatment Lyme disease symptoms score: Developing a new tool for research // PLoS One. 2019. Vol. 14, N 11. P. e0225012. doi: 10.1371/journal.pone.0225012
21. Миноранская Н.С., Усков А.Н., Сарап П.В. Значение иммунного статуса для прогноза хронизации боррелиозной инфекции // Журнал инфектологии. 2014. Т. 6, № 1. С. 35–40.
22. Суздальцев А.А., Каравашкин Н.В., Кулагина А.П. Клинико-эпидемиологические аспекты иксодового клещевого боррелиоза в Самарской области // Медицинский вестник Башкортостана. 2021. Т. 16, № 3 (93). С. 27–32.
23. Бондаренко А.Л., Сапожникова В.В. Анализ клинико-эпидемиологических, лабораторных показателей и цитокинового статуса у пациентов с эритемной и безэритемной формами иксодового клещевого боррелиоза // Инфекционные болезни. 2018. Т. 16, № 2. С. 34–42. doi: 10.20953/1729-9225-2018-2-34-42
24. Blom K., Cuapio A., Sandberg J.T., et al. Cell-mediated immune responses and immunopathogenesis of human tick-borne encephalitis virus-infection // Front Immunol. 2018. Vol. 9. P. 2174. doi: 10.3389/fimmu.2018.02174
25. Rutkowska E., Kwiecień I., Kulik K., et al. Usefulness of the new hematological parameter: reactive lymphocytes RE-LYMP with flow cytometry markers of inflammation in COVID-19 // Cells. 2021. Vol. 10, N 1. P. 82. doi: 10.3390/cells10010082
26. Georgakopoulou V.E., Makrodimetri S., Triantafyllou M., et al. Immature granulocytes: Innovative biomarker for SARS-CoV-2 infection // Mol Med Rep. 2022. Vol. 26, N 1. P. 217. doi: 10.3892/mmr.2022.12733
27. Rizo-Télez S.A., Méndez-García L.A., Flores-Rebollo C., et al. The neutrophil-to-monocyte ratio and lymphocyte-to-neutrophil ratio at admission predict in-hospital mortality in Mexican patients with severe SARS-CoV-2 Infection (COVID-19) // Microorganisms. 2020. Vol. 8, N 10. P. 1560. doi: 10.3390/microorganisms8101560
28. Oehadian A., Michels M., de Mast Q., et al. New parameters available on Sysmex XE-5000 hematology analyzers contribute to differentiating dengue from leptospirosis and enteric fever // Int J Lab Hematol. 2015. Vol. 37, N 6. P. 861–868. doi: 10.1111/ijlh.12422

REFERENCES

1. Lobzin YuV, Uskov AN, Kozlov SS. *Laim-borrelioz: iksodovyye kleshchevyye borreliozy*. Saint Petersburg: Foliant; 2000. 156 p. (In Russ).
2. Ierusalimsky AP. *Kleshchevoi entsefalit: rukovodstvo dlya vrachei*. Novosibirsk: Gosudarstvennaya meditsinskaya akademiya MZ RF; 2001. 360 p. (In Russ).

3. Bondarenko AL, Zykova IV, Abbasova SV, Tikhomolova EG, Nekhoroshkina EL. Mixed infection of tick-borne encephalitis and ixodes tick-borne borrelioses. *Infektsionnye Bolezni*. 2011;9(4):54–63. (In Russ).
4. Subbotin AV, Semenov VA, Etenko DA. Problema sovremennykh smeshannykh neiroinfektsii, peredayushchikhsya iksodovymi kleshchami. *The Russian Archives of Internal Medicine*. 2012; (2(4)):35–39. (In Russ).
5. Konkova-Reidman AB, Zlobin VI. Clinical polymorphism of ixodes tick-borne borrelioses (mixed infection with tick-borne encephalitis) on the territory of South-Ural region of Russia. *Sibirskii meditsinskii zhurnal (Irkutsk)*. 2011;100(1):17–20. (In Russ).
6. Minoranskaya NS, Minoranskaya EI. Clinical and epidemiologic characteristics of Lyme borreliosis and tick-borne encephalitis mixed infection in Krasnoyarsk kray. *Kazan Medical Journal*. 2013; 94(2):211–215. (In Russ).
7. Andronova NV, Minoranskaya NS, Minoranskaya EI. The specific immune response and some remote results in the acute course of tick-borne borreliosis and mixed-infection of tick-borne encephalitis and tick-borne borreliosis. *Sibirskii meditsinskii zhurnal (Irkutsk)*. 2011;100(1):54–57. (In Russ).
8. Tanner L, Schreiber M, Low JG, et al. Decision tree algorithms predict the diagnosis and outcome of dengue fever in the early phase of illness. *PLoS Negl Trop Dis*. 2008;2(3):e196. doi: 10.1371/journal.pntd.0000196
9. Tamibmaniam J, Hussin N, Cheah WK, Ng KS, Muninathan P. Proposal of a clinical decision tree algorithm using factors associated with severe dengue infection. *PLoS One*. 2016;11(8):e0161696. doi: 10.1371/journal.pone.0161696
10. Dobrijević D, Andrijević L, Antić J, Rakić G, Pastor K. Hemogram-based decision tree models for discriminating COVID-19 from RSV in infants. *J Clin Lab Anal*. 2023;37(6):e24862. doi: 10.1002/jcla.24862
11. Kumaran M, Pham TM, Wang K, et al. Predicting the risk factors associated with severe outcomes among COVID-19 patients—decision tree modeling approach. *Front Public Health*. 2022;10:838514. doi: 10.3389/fpubh.2022.838514
12. Bruce P, Bruce A. *Practical Statistics for Data Scientists*. Saint Petersburg: BKhV-Peterburg; 2018. 304 p. (In Russ).
13. Kowalska-Kępczyńska A, Mleczko M, Domerecka W, Krasowska D, Donica H. Assessment of immune cell activation in pemphigus. *Cells*. 2022;11(12):1912. doi: 10.3390/cells11121912
14. Grebennikova IV, Lidokhova OV, Makeeva AV, et al. Age-dependend changes of leukocyte indices in COVID-19. *Naučno-medicinskij vestnik Central'nogo Černozem'â*. 2022;(87): 9–15. (In Russ).
15. Petrie A, Sabin K. *Medical statistics at a glance*. Leonov VP, editor. Moscow: GEOTAR-Media; 2015. 216 p. (In Russ).
16. Huang Y, Li S, Lin B, et al. Early detection of college students' psychological problems based on decision tree model. *Front Psychol*. 2022;13:946998. doi: 10.3389/fpsyg.2022.946998
17. Mukhopadhyay S. *Advanced data analytics using Python: with machine learning, deep learning and NLP examples*. Kolkata : Apress Berkeley; 2018. 186 p. doi: 10.1007/978-1-4842-3450-1
18. Kalafi EY, Nor NAM, Taib NA, et al. Machine learning and deep learning approaches in breast cancer survival prediction using clinical data. *Folia Biol (Praha)*. 2019;65(5-6):212–220.
19. Lakshmanan V, Sara Robinson S, Munn M. *Machine learning. Design patterns*. Saint Petersburg: BKhV-Peterburg; 2022. 448 p. (In Russ).
20. Turk SP, Lumbard K, Liepshutz K, et al. Post-treatment Lyme disease symptoms score: Developing a new tool for research. *PLoS One*. 2019;14(11):e0225012. doi: 10.1371/journal.pone.0225012
21. Minoranskaya NS, Uskov AN, Sarap PV. Importance of immune status for prognosis chronic borreliosis infections. *Jurnal infektologii*. 2014;6(1): 35–40. (In Russ).
22. Suzdaltcev AA, Karavashkin NV, Kulagina AP. Clinical and epidemiological aspects of ixodic tick-borne borreliosis in the Samara region. *Meditsinskii vestnik Bashkortostana*. 2021;16(3(93)):27–32. (In Russ).
23. Bondarenko AL, Sapozhnikova VV. Analysis of clinical-epidemiological, laboratory parameters and cytokine status in patients with erythematous and non-erythematous forms of ixodes tick borreliosis. *Infektsionnye Bolezni*. 2018;16(2):34–42. (In Russ). doi: 10.20953/1729-9225-2018-2-34-42
24. Blom K, Cuapio A, Sandberg JT, et al. Cell-mediated immune responses and immunopathogenesis of human tick-borne encephalitis virus-infection. *Front Immunol*. 2018;9:2174. doi: 10.3389/fimmu.2018.02174
25. Rutkowska E, Kwiecień I, Kulik K, et al. Usefulness of the new hematological parameter: reactive lymphocytes RE-LYMP with flow cytometry markers of inflammation in COVID-19. *Cells*. 2021;10(1):82. doi: 10.3390/cells10010082
26. Georgakopoulou VE, Makrodimitri S, Triantafyllou M, et al. Immature granulocytes: Innovative biomarker for SARS-CoV-2 infection. *Mol Med Rep*. 2022;26(1):217. doi: 10.3892/mmr.2022.12733
27. Rizo-Téllez SA, Méndez-García LA, Flores-Rebollo C, et al. The neutrophil-to-monocyte ratio and lymphocyte-to-neutrophil ratio at admission predict in-hospital mortality in Mexican patients with severe SARS-CoV-2 Infection (COVID-19). *Microorganisms*. 2020;8(10):1560. doi: 10.3390/microorganisms8101560
28. Oehadian A, Michels M, de Mast Q, et al. New parameters available on Sysmex XE-5000 hematology analyzers contribute to differentiating dengue from leptospirosis and enteric fever. *Int J Lab Hematol*. 2015;37(6):861–868. doi: 10.1111/ijlh.12422

ОБ АВТОРАХ

* **Ильинских Екатерина Николаевна**, д.м.н., доцент;
адрес: Россия, 634050, Томск, Московский тракт, д. 2;
ORCID: 0000-0001-7646-6905;
eLibrary SPIN: 5245-5958;
e-mail: infconf2009@mail.ru

AUTHORS' INFO

* **Ekaterina N. Ilyinskikh**, MD, Dr. Sci. (Med.), Associate Professor;
address: 2 Moskovsky trakt, 634050 Tomsk, Russia;
ORCID: 0000-0001-7646-6905;
eLibrary SPIN: 5245-5958;
e-mail: infconf2009@mail.ru

Филатова Евгения Николаевна;

ORCID: 0000-0001-9951-8632;
eLibrary SPIN: 8094-3417;
e-mail: synamber@mail.ru

Самойлов Кирилл Владимирович;

ORCID: 0000-0002-8477-8551;
eLibrary SPIN: 4710-0894;
e-mail: samoilov.krl@gmail.com

Семёнова Алина Васильевна;

ORCID: 0000-0001-5195-3897;
eLibrary SPIN: 2690-1166;
e-mail: wind_of_change95@mail.ru

Аксёнов Сергей Владимирович, к.т.н., доцент;

ORCID: 0000-0002-1251-7133;
eLibrary SPIN: 2229-4552;
e-mail: axyonov@tpu.ru

Evgenia N. Filatova, MD;

ORCID: 0000-0001-9951-8632;
eLibrary SPIN: 8094-3417;
e-mail: synamber@mail.ru

Kirill V. Samoylov, MD;

ORCID: 0000-0002-8477-8551;
eLibrary SPIN: 4710-0894;
e-mail: samoilov.krl@gmail.com

Alina V. Semenova, MD;

ORCID: 0000-0001-5195-3897;
eLibrary SPIN: 2690-1166;
e-mail: wind_of_change95@mail.ru

Sergey V. Axyonov, Cand. Sci. (Eng.), Associate Professor;

ORCID: 0000-0002-1251-7133;
eLibrary SPIN: 2229-4552;
e-mail: axyonov@tpu.ru

* Автор, ответственный за переписку / Corresponding author