

Факторы риска тяжелого и затяжного течения острых респираторных вирусных инфекций у пациентов молодого возраста

Рычкова О.А.¹,
Кашуба Э.А.¹,
Рукина Н.Ю.²,
Перминов Н.В.³,
Сизов Д.А.⁴

¹ Федеральное государственное бюджетное образовательное учреждение высшего образования «Тюменский государственный медицинский университет» Министерства здравоохранения Российской Федерации, 625023, г. Тюмень, Российская Федерация

² Федеральное государственное бюджетное образовательное учреждение высшего образования «Дальневосточный государственный медицинский университет» Министерства здравоохранения Российской Федерации, 68000, г. Хабаровск, Российская Федерация

³ Федеральное государственное казенное учреждение «301 военный клинический госпиталь» Министерства обороны Российской Федерации, 68000, г. Хабаровск, Российская Федерация

⁴ Государственное бюджетное учреждение здравоохранения города Салехард «Окружная клиническая больница Департамента здравоохранения города Салехард», 629000, г. Салехард, Российская Федерация

Острые респираторные инфекции остаются одними из наиболее распространенных инфекционных заболеваний, которые в основном протекают в легкой форме, но есть вероятность развития тяжелых форм заболевания. В статье рассмотрены возможности раннего прогнозирования тяжелого течения острой респираторной вирусной инфекции (ОРВИ) у пациентов молодого возраста.

Цель исследования – разработка математической прогностической модели длительности течения и степени тяжести ОРВИ для использования на этапах госпитализации пациентов молодого возраста.

Материал и методы. Проведено исследование данных 120 людей (90+30) молодого возраста (от 18 до 30 лет), из которых 90 пациентов, проходивших лечение в инфекционном отделении ФГКУ «301 ВКГ» Минобороны России (Хабаровск) с диагнозом «острая респираторная вирусная инфекция», и 30 условно-здоровых молодых людей. Диагноз «острая респираторная вирусная инфекция» был установлен на основании эпидемиологических, клинических и лабораторных данных. Этиология болезни установлена методом полимеразной цепной реакции.

Результаты и обсуждение. Статистическая обработка полученных данных показала, что из 95 оцениваемых клинико-лабораторных показателей только 15 были определены как существенно влияющие на длительность и тяжесть течения болезни ($p < 0,001$). Для отобранных показателей были рассчитаны диагностические значения и построена математическая модель прогнозирования длительности и тяжести течения болезни, имеющая высокий уровень достоверности, – 95%.

Заключение. ROC-анализ построенных «деревьев решения» показал их высокое прогностическое значение. Прогностическое качество построенной модели имеет средний уровень. Значение $AuROC=0,84$ позволяет считать, что смоделированное «дерево решений» имеет среднее прогностическое значение. Диагностическая ценность отобранных показателей – средняя ($AUC=0,84$); чувствительность и специфичность – высокие (88,2 и 69,5% соответственно). Такое сочетание чувствительности и специфичности модели логистической регрессии позволяет рекомендовать использование выбранных показателей на ранних этапах прогнозирования длительности и тяжести течения ОРВИ.

Финансирование. Исследование не имело финансовой поддержки.

Конфликт интересов. Авторы сообщают об отсутствии конфликта интересов.

Ключевые слова:

острая респираторная вирусная инфекция; взрослые пациенты (мужчины); логистическая регрессия; дерево принятия решений

Вклад авторов. Концепция и дизайн исследования – Сизов Д.А., Рукина Н.Ю., Перминов Н.В., Рычкова О.А., Кашуба Э.А.; сбор материала – Сизов Д.А., Перминов Н.В.; обработка материала, статистическая обработка данных, написание текста статьи – Сизов Д.А., Рукина Н.Ю., Перминов Н.В., Рычкова О.А., Кашуба Э.А.; редактирование – Кашуба Э.А., Рычкова О.А.; утверждение окончательного варианта статьи – Рычкова О.А., Кашуба Э.А.; ответственность за целостность всех частей статьи – Сизов Д.А., Рычкова О.А.

Для цитирования: Рычкова О.А., Кашуба Э.А., Рукина Н.Ю., Перминов Н.В., Сизов Д.А. Факторы риска тяжелого и затяжного течения острых респираторных вирусных инфекций у пациентов молодого возраста // Инфекционные болезни: новости, мнения, обучение. 2022. Т. 11, № 4. С. 56–63. DOI: <https://doi.org/10.33029/2305-3496-2022-11-4-56-63>

Статья поступила в редакцию 27.02.2022. **Принята в печать** 24.10.2022.

Risk factors for severe and prolonged acute respiratory viral infections in young patients

*Rychkova O.A.*¹, *Kashuba E.A.*¹,
*Rukina N.Yu.*², *Perminov N.V.*³,
*Sizov D.A.*⁴

¹Tyumen State Medical University of the Ministry of Health of the Russian Federation, 625023, Tyumen, Russian Federation
²Far Eastern State Medical University of the Ministry of Health of the Russian Federation, 68000, Khabarovsk, Russian Federation
³301 Military Clinical Hospital of the Ministry of Defense of the Russian Federation, 68000, Khabarovsk, Russian Federation
⁴District Clinical Hospital of the Department of Health of the City of Salekhard, 629000, Salekhard, Russian Federation

Acute respiratory infections remain one of the most common infectious diseases, which are mostly mild, and there is also a possibility of developing severe forms of the disease. The article considered the possibilities of early prediction of severe course of acute respiratory viral infection in young patients.

The aim – creation of a mathematical prognostic model of the duration of the course and severity of acute respiratory viral infection for use at the stages of hospitalization of young patients.

Material and methods. A study was conducted on 120 people (90+30) of young age (from 18 to 30 years old), of which 90 people were treated in the Infectious Diseases Department of the 301 VKG of the Ministry of Defense of the Russian Federation of Khabarovsk with a diagnosis of “acute respiratory viral infection” and 30 – conditionally healthy young people. The diagnosis of acute respiratory viral infection was established on the basis of epidemiological, clinical and laboratory data. The etiology of the disease was established by PCR.

Results and discussion. Statistical processing of the data obtained showed that out of 95 assessed clinical and laboratory parameters, only 15 were identified as significantly affecting the duration and severity of the disease ($p < 0.001$). For the selected indicators, diagnostic values were calculated and a mathematical model was built to predict the duration and severity of the course of the disease, which has a high level of reliability – 95%.

Conclusion. ROC analysis of the constructed “decision trees” showed their high prognostic value. The predictive quality of the constructed model has an average level. An AuROC value of 0.84 suggests that the simulated decision tree has an average predictive value. The diagnostic value of the selected indicators is average (AUC=0.84); sensitivity – high (88.2%), specificity (69.5%) – high. This combination of sensitivity and specificity of the logistic regression model allows us to recommend the use of the selected indicators in the early stages of predicting the duration and severity of an acute respiratory viral infection.

Funding. The study was not funded.

Conflict of interest. The authors report no conflict of interest.

Contribution. The concept and design of the study – Sizov D.A., Rukina N.Yu., Perminov N.V., Rychkova O.A., Kashuba E.A.; collection of material – Sizov D.A., Perminov N.V.; material processing, statistical data processing, writing the text of the article – Sizov D.A., Rukina N.Yu., Perminov N.V., Rychkova O.A., Kashuba E.A.; editing – Kashuba E.A., Rychkova O.A.; approval of the final version of the article – Rychkova O.A., Kashuba E.A.; responsibility for the integrity of all parts of the article – Sizov D.A., Rychkova O.A.

For citation: Rychkova O.A., Kashuba E.A., Rukina N.Yu., Perminov N.V., Sizov D.A. Risk factors for severe and prolonged acute respiratory viral infections in young patients. *Infektsionnye bolezni: novosti, mneniya, obuchenie* [Infectious Diseases: News, Opinions, Training]. 2022; 11 (4): 56–63. DOI: <https://doi.org/10.33029/2305-3496-2022-11-4-56-63>

Received 27.02.2022. **Accepted** 24.10.2022.

Keywords:

acute respiratory viral infection; adult patients (male); logistic regression; decision tree

Заболлеваемость острыми инфекциями верхних дыхательных путей неуточненной этиологии и гриппом в России, по данным официальной статистики, составляет 19–20 тыс. на 100 тыс. населения ежегодно, превы-

шая в десятки раз аналогичные показатели для других инфекционных болезней, и составляет 90% всех регистрируемых случаев инфекционных и паразитарных болезней [1]. Большинство пациентов обращаются к терапевтам по

Таблица 1. Клинико-лабораторные критерии включенных в исследование молодых людей с острой респираторной вирусной инфекцией

Показатель	Характеристика показателя
<i>Клинические критерии</i>	
День болезни	От 1-го до 5-го
Температура тела	Нормальная, субфебрильная, фебрильная
Фарингит	Наличие или отсутствие (0/1)
Насморк (0/1/2/3)	Степень выраженности от 1 до 3
Гипертрофия миндалин (0/1/2/3)	Степень выраженности от 1 до 3
Регионарный лимфаденит (0/1)	Наличие или отсутствие (0/1)
Боль при движении глазами (0/1)	Наличие или отсутствие (0/1)
<i>Лабораторные критерии (гематологические)</i>	
Лейкоциты (WBC)	$4-9 \times 10^9/\text{л}$
Эритроциты (RBC)	$3,8-5,6 \times 10^{12}/\text{л}$
Гемоглобин (HGB)	120–170 г/л
Гематокрит (HCT)	34,0–48,0%
Тромбоциты (PLT)	$150-420 \times 10^9/\text{л}$
Тромбокрит (PCT)	150–400%
Средний объем эритроцита (MCV)	80–100 мкм ³
Эритроцитарный индекс (MCH)	26,0–34,0 пг
Насыщение эритроцитов Hb (MCHC)	300–360 г/л
Ширина эритроцитов (RDW)	10,0–15,0%
Средний объем тромбоцитов (MPV)	7,0–10,0 мкм ³
Распределение тромбоцитов по V (PDW)	10,0–18,0%
Лимфоциты (LYM)	19,0–45,0%
Моноциты (MON)	2,0–11,0%
Нейтрофилы (GRA)	45,0–80,0%
Лимфоциты, абс.	$1,2-3,5 \times 10^9/\text{л}$
Моноциты, абс.	$0,1-0,6 \times 10^9/\text{л}$
Нейтрофилы, абс.	$1,6-6,9 \times 10^9/\text{л}$
Нейтрофилы сегментоядерные	47–67%
Нейтрофилы палочкоядерные	1–6%
Эозинофилы	0,5–5%
Базофилы	0–1%
Плазматические клетки	1–2%
Токсическая зернистость	В норме отсутствует
Скорость оседания эритроцитов	м – 1–10 мм/ч
<i>Лабораторные критерии (иммунологические)</i>	
MCP-1 (CCL2)	15–50 нг/мл
Интерлейкин 8	2–60 пг/мл
Интерлейкин 18	31,25–60 пг/мл
Интерферон γ	1–128 ед/мл
STAT-1	0,156–10 нг/мл

Примечание. STAT-1 – представитель семейства транскрипционных факторов преобразователей сигналов и активаторов транскрипции. STAT-1 участвует в положительной регуляции генов по сигналам интерферона типа I, типа II или типа III.

поводу острой респираторной инфекции [2]. На инфекции верхних дыхательных путей приходится около 10 млн амбулаторных посещений в год. Инфекции верхних дыхательных путей являются причиной более 20 млн пропущенных дней в школе и более 20 млн дней пропущенной работы, что создает большое экономическое бремя [3]. Тяжелая респираторная инфекция является основной причиной госпитализации в отделение интенсивной терапии пациентов с ослабленным иммунитетом, которые подвержены риску гипоксемической острой дыхательной недостаточности и сепсису [4]. Тяжелая гриппозная инфекция сопровождается нарушением регуляции цитокинов и хемокинового ответа [5].

Инфекционные болезни являются результатом дисгармоничного экологического взаимодействия микробного инфекционного агента (бактерий, грибов, паразитов или вирусов, за исключением прионов, которые являются инфекционными белками) с хозяином, причем динамика этого взаимодействия подвергается модулирующему влиянию нескольких факторов, таких как окружающая среда, биологические и социальные факторы [6].

Разработка и использование математических моделей развития эпидемического процесса широко признаны в различных областях, таких как экология или эпидемиология [7]. Традиционная математическая эпидемиология как метод количественного исследования широко исполь-

зовалась в области распространения эпидемий. Двумя типичными моделями распространения эпидемии являются восприимчивые-инфекционные восстановленные (SIR) [8] и восприимчивые-инфекционные восприимчивые (SIS) [9]. Эти модели используются для различных целей: понимания динамики распространения инфекционных заболеваний, информирования о политике или руководящих принципах общественного здравоохранения, например, с помощью моделирования возможных дополнительных вмешательств [10]. В случае чрезвычайных ситуаций в области общественного здравоохранения одна из основных задач заключается в построении адекватной модели распространения эпидемии. Она может дать обоснованное объяснение того, какие факторы могут повлиять на распространение возбудителя, и своевременно сформулировать эффективные меры профилактики и контроля. Поэтому очень важно использовать метод динамической модели для оценки передачи инфекционных болезней [11].

Цель исследования – разработка математической прогностической модели длительности течения и степени тяжести острой респираторной вирусной инфекции (ОРВИ) для использования на этапах госпитализации пациентов молодого возраста.

Материал и методы

Проведено обследование 120 пациентов (мужчин молодого возраста от 18 до 30 лет), из которых 90 проходили лечение в инфекционном отделении ФГКУ «301 ВКГ» Минобороны России с диагнозом «ОРВИ». Контрольная группа – 30 условно-здоровых молодых людей.

Критерии включения: возраст от 18 до 30 лет, отсутствие коморбидной патологии, день заболевания – от 1-го до 5-го, отсутствие роста патогенных микроорганизмов при проведении бактериологического исследования. Для поиска информативных показателей было изучено 95 признаков, в том числе лабораторные и эпидемиологические, с учетом возраста пациентов, длительности лихорадки, иммунологических показателей (табл. 1).

Для установления прогностически значимых факторов риска, влияющих на степень тяжести ОРВИ, всем пациентам был проведен физикальный осмотр и расширенный лабораторный спектр исследований.

Все участники исследования подписали форму информированного добровольного согласия, утвержденную локальным этическим комитетом ФГБОУ ВО ДВГМУ Минздрава России (протокол № 7 от 10.04.2021). Разрешение на исследование получено в локальном этическом комитете. Больным с острой респираторной инфекцией проводили стандартное клиническое и лабораторное обследование.

Этиологическую диагностику ОРВИ осуществляли методом полимеразной цепной реакции (ПЦР) с гибридационно-флюоресцентной детекцией «АмплиСенс® ОРВИ-скрин-FL». Определяли РНК респираторно-синцитиального вируса (human Respiratory Syncytial virus – hRSv), метапневмовируса (human Metapneumovirus-hMPv), вирусов парагриппа 1, 2, 3 и 4-го типов (human Parainfluenza virus-1-4 – hPIv), коронавируса (human Coronavirus – hCov), риновиру-

сов (human Rhinovirus – hRv), ДНК аденовирусов групп В, С и Е (human Adenovirus В, С, Е – hAdv) и бокавируса (human Bocavirus – hBov) в клиническом материале.

Бактериологические методы: посевы проводили на среды, наиболее благоприятные для культивирования кокков, гемофильной палочки – мясопептонный агар с добавлением 5% крови кролика и агар Левинтала. Забор биологического материала осуществляли в транспортную среду Stuart. У 90 пациентов был проведен ПЦР-анализ – смывы из ротоглотки, посевы из зева и носа на микрофлору с расшифровкой условно-патогенной флоры.

Образцы периферической крови в объеме 2 мл брали у больных при госпитализации и у практически здоровых людей. Показатели гемограммы исследовали на анализаторе Sysmex XP-300. Для определения концентрации интерлейкинов количественным методом использовали автоматический EVOLIS Twin Plus (Bio-Rad, Франция). Исследование сывороток крови проводили методом иммуноферментного анализа (ИФА) с использованием наборов реагентов: ИФА набор Human Interleukin 18, IL-18, 96; Human Signal transducer and activator of transcription 1 (STAT1) ELISA Kit, 96T фирмы CUSABIO (USA); ИФА набор MCP-1 (монокитарный хемотаксический белок-1), 96; ИФА набор Human Interleukin 8, IL-8, 96, и ИФА гамма-интерферон («Вектор Бест», Новосибирск), на иммуноферментном автоматическом анализаторе EVOLIS Twin Plus (Bio-Rad, Франция), следуя техническим рекомендациям.

Статистическую обработку данных осуществляли с помощью пакетов прикладных программ Statistica 10 и SAS JMP 11. Все непараметрические показатели переведены в балльную систему. Результаты модели логистической регрессии представлены как скорректированные коэффициенты отношения шансов с учетом 95% доверительных интервалов (ДИ), также был определен относительный риск.

Сравнивали показатели по *t*-критерию Стьюдента для независимых выборок и критерия χ^2 для оценки категориальных различий. Для оценки влияния каждой независимой переменной на интересующие результаты, переменные, достигающие в двумерном анализе значения $p \leq 0,25$, были включены в многофакторную модель логистической регрессии. Проведен однофакторный анализ прогнозирования длительности заболевания с учетом количественных и бинарных факторов. Статистическую значимость влияния оцениваемого фактора на бинарную целевую переменную определяли с помощью критерия χ^2 Пирсона. Все факторы сортировали по убыванию значимости (величины χ^2), и, таким образом, были отобраны ключевые факторы, влияющие на увеличение длительности заболевания [12, 13]. В исходных данных ввели целевую переменную «длительность заболевания» (когда пациент болеет 5 дней и больше).

Анализ зависимости переменных факторов при сопоставлении 2 временных периодов осуществляли на основе непараметрического критерия Фридмана.

Для моделирования бинарных показателей использованы «деревья решения». Данный метод позволяет последовательно изучать эффект влияния отдельных переменных, а также предоставляет возможность выполнять «одномерное ветвление» для анализа вклада отдельных показателей

Таблица 2. Ключевые клинико-лабораторные показатели, отобранные с применением метода математического моделирования

Фактор (показатель)	Длительность заболевания (частота), %		Изменение риска (95% ДИ)	Относительный риск (95% ДИ)	Достовер- ность (p)	χ^2
	наличие фактора	отсутствие фактора				
<i>Клинические критерии</i>						
Длительность болезни ≥ 5 дней	4 (6,3)	15 (62,5)	56,2 (35,9–76,4)	9,84 (3,63–26,70)	<0,0001	32,10
Наличие фарингита на 5-й день болезни	1 (1,9)	18 (51,4)	49,5 (32,5–65,5)	26,74 (3,74–91,29)	<0,0001	30,03
Сохранение температуры тела $\geq 37,0$ °C на протяжении 5 дней	2 (3,7)	17 (51,5)	47,8 (30,0–65,6)	13,91 (3,43–56,39)	<0,0001	27,43
Сохранение температуры тела $\geq 38,0$ °C на протяжении 5 дней	10 (15,9)	9 (37,5)	21,6 (0,3–43,0)	2,36 (1,10–5,09)	0,0291	4,76
Сохранение насморка на 5-й день заболевания	5 (9,3)	14 (42,4)	33,2 (14,6–51,7)	4,58 (1,82–11,55)	0,0003	13,19
Максимальная температура тела $\geq 38,0$ °C	8 (13,6)	11 (39,3)	25,7 (5,6–45,8)	2,9 (1,31–6,40)	0,0067	7,36
Лихорадка, температура тела в день обращения $\geq 38,0$ °C	7 (12,1)	10 (38,5)	26,4 (5,9–46,9)	3,19 (1,36–7,44)	0,0054	7,74
<i>Лабораторные критерии</i>						
Нейтрофилы <5,0, тыс./мкл	1 (3,1)	18 (32,7)	29,6 (15,8–43,4)	10,47 (1,47–74,79)	0,0013	10,38
Средний объем тромбоцитов $\geq 8,8$ мкм ³	4 (9,1)	15 (34,9)	25,8 (9,2–42,4)	3,84 (1,38–10,64)	0,0036	8,47
Уровень ИЛ-8 $\geq 9,3$ пг/мл	4 (9,3)	14 (33,3)	24,0 (7,3–40,7)	3,58 (1,28–10,00)	0,0067	7,35
Лимфоциты <1,3 $\times 10^9$ /л	7 (13,7)	12 (33,3)	19,6 (1,5–37,7)	2,43 (1,06–5,56)	0,0292	4,75
Моноциты $\geq 3,5 \times 10^9$ /л	1 (4,5)	18 (27,7)	23,1 (9,2–37,1)	6,09 (0,86–43,03)	0,0231	5,15
Лейкоциты <6,2 $\times 10^9$ /л	5 (11,6)	14 (31,8)	20,2 (3,4–37,0)	2,74 (1,08–6,94)	0,0227	5,19
Эритроциты $\geq 5,1 \times 10^{12}$ /л	8 (14,0)	11 (36,7)	22,6 (3,2–42,1)	2,61 (1,18–5,79)	0,0152	5,89
Гемоглобин ≥ 153 г/л	9 (15,0)	10 (37,0)	22,0 (1,7–42,4)	2,47 (1,13–5,37)	0,0214	5,29
Тромбоциты <171,0 $\times 10^9$ /л	9 (15,0)	10 (37,0)	22,0 (1,7–42,4)	2,47 (1,13–5,37)	0,0214	5,29

и работать с показателями разных типов (количественных и номинальных), позволяет дать интуитивно понятную интерпретацию каждого рискованного класса с оценкой риска и размера [14].

Для анализа качества смоделированных «деревьев» использован ROC-анализ. Уровень статистической значимости различий был зафиксирован на уровне 0,05.

Результаты и обсуждение

Средний возраст обследованных пациентов составил $20 \pm 0,89$ года. Во всех исследованных 90 образцах были обнаружены вирусы группы ОРВИ. В структуре выявленных нозологических форм ОРВИ преобладали риновирусная 40 (44,4%) и гриппозная инфекции – 25 (27,8%), парагриппозная и респираторно-синцитиальная инфекции составили 14 (15,6%) и 11 (12,2%) соответственно.

С помощью метода математического моделирования определен относительный риск и рассчитана вероятность длительности заболевания (табл. 2). Для построения таблицы были взяты следующие параметры: клинические (длительность заболевания, лихорадка, фарингит, насморк), лабораторные (уровень тромбоцитов, лейкоцитов, нейтрофилов, ИЛ-8 и ИЛ-18, уровень моноцитарного хемотаксического белка, преобразователя сигнала и активатора транскрипции 1).

По результатам проведенного однофакторного анализа рисков увеличения длительности заболевания выявлено, что ключевыми статистически значимыми среди них являются: длительность болезни более 5 дней, клинические проявления фарингита, повышение температуры тела более субфебрильной $\geq 37,0$ °C (уровень риска более 51,4%). Наличие одного из этих 3 ключевых прогностических факторов свидетельствует, что заболевание будет протекать длительно, с осложненным течением с вероятностью в 0,98. Методом прямого пошагового дискриминантного анализа из 95 исследуемых клинико-лабораторных показателей были выявлены 15, которые достоверно различались в изучаемых группах и имели статистическую значимость при оценке влияния на длительность заболевания. При наличии у пациентов указанных факторов риск пролонгирования срока болезни составлял от 27,5 до 100,0%. Ведущими статистически значимыми факторами с уровнем относительных рисков от 51,4 до 62,5% являются продолжительность болезни ≥ 5 дней, присутствие клинических проявлений фарингита на 5-й день и сохранение температуры тела $\geq 37,0$ °C на протяжении 5 дней. Наименьшие значения показателя относительного риска получены в отношении лабораторных признаков – уровень лимфоцитов <1,3 $\times 10^9$ /л, уровень моноцитов $\geq 3,5 \times 10^9$ /л, которые повышают уровень относительного риска от 27,7 до 37,5%.

Длительность болезни – период, который начинается с момента возникновения заболевания и продолжается

до полного выздоровления. Длительное заболевание является более узким критерием, это превышение сроков среднего количества дней, определенных при конкретном заболевании, т.е. больше средних сроков течения болезни. Например, при заболеваниях респираторными вирусными инфекциями средний срок составляет от 3 до 5 дней госпитализации в профильных стационарах. В рамках применения метода многомерного анализа всех пациентов разделили на несколько рискованных классов целевого события (например, длительность болезни) с помощью комбинации влияющих факторов и далее ранжировали классы по уровню риска. Одно из построенных «деревьев классификации» представляет собой довольно простое «дерево решений» с несколькими вопросами. Это бинарное «дерево классификации» можно описать следующей фразой: «Если при госпитализации у пациента уровень среднего объема тромбоцитов $\geq 8,8$ мкм³ и его возраст более 20 лет, то в этом и только в этом случае следует ожидать, что заболевание будет протекать более 5 дней, с осложненным течением с риском равным 45,2%!». Из этого предложения несложно представить себе соответствующее «дерево решений». Вопросы задаются последовательно (иерархически), и окончательное решение зависит от ответов на все предыдущие вопросы [16].

«Деревья классификации» – метод, позволяющий прогнозировать риски целевых событий в зависимости от соответствующих уровней независимых факторов, а также разделять респондентов на рискованные классы [17]. Основная задача «деревя решений» – выделение групп пациентов на основании нескольких факторов (в отличие от анализа риска, где каждый фактор рассматривается отдельно) и вычисление относительного риска наступления негативного события при той или другой комбинации факторов. Негативное событие – это результат, при котором модель правильно прогнозирует отрицательный класс.

Каждое «ветвление дерева» происходит по конкретному фактору, разделяя пациентов с наличием фактора и тех, у которых он отсутствует. Таким образом, пациенты были разделены на несколько групп (классов), в каждом из которых имеется свой уровень относительного риска. Рисковый класс – группа пациентов с наличием фактора относительного риска. Для удобства восприятия отсортированные рискованные классы пациентов внесены в таблицу упорядочивания от максимального риска до минимального. Данные представлены на рис. 1.

Всего с помощью «деревя решений» было выделено 4 рискованных класса (величина риска длительности болезни более 5 дней). Наиболее высокий риск (риск = 45,2%, $n=31$) развития длительности заболевания (более 5 дней) возможен

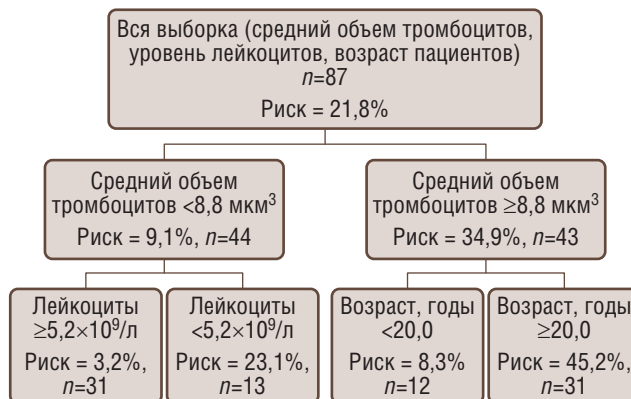


Рис. 1. «Дерево классификации» длительности болезни (факторов, которые влияют на ее продолжительность) – более 5 дней. Уровень риска всей выборки при комбинации 3 факторов (средний объем тромбоцитов, уровень лейкоцитов и возраст)

у пациентов со следующей комбинацией факторов: средний объем тромбоцитов $\geq 8,8$ мкм³, возраст $\geq 20,0$ года. Наименьший уровень риска (риск = 3,2%, $n=31$) развития длительности заболевания (более 5 дней) возможен для следующей комбинации факторов: средний объем тромбоцитов $< 8,8$ мкм³ и уровень лейкоцитов $\geq 5,2 \times 10^9$ /л. Самым большим классом с $n=31$ наблюдений и уровнем риска в 3,2% является следующая комбинация факторов: средний объем тромбоцитов $\geq 8,8$ мкм³ и возраст $\geq 20,0$ года. У 13 пациентов уровень риска находился на среднем уровне.

Определены ключевые факторы, которые могли оказывать влияние на частоту увеличения длительности болезни более 5 дней (табл. 3).

На рис. 2 и в табл. 4 приведены результаты ROC-анализа и прогнозные показатели построенного «деревя решений» с учетом длительности заболевания. Точка отсечения представляет собой оптимальную границу отделения положительного прогноза от отрицательного.

Прогнозное качество построенной модели имеет средний уровень.

Заключение

Персонализированная медицина – совокупность методов профилактики патологического состояния, диагностики и лечения в случае его возникновения, основанных на индивидуальных особенностях пациента. Этот подход также учитывает убеждения каждого пациента в отношении здоровья,

Таблица 3. Отсортированные рискованные классы пациентов для длительности заболевания по убыванию риска

Показатель рискованного класса	Объем класса, n	Доля класса в целом по выборке, %	Риск, %
Средний объем тромбоцитов $\geq 8,8$ мкм ³ ; возраст $\geq 20,0$ года	31	35,6	45,2
Средний объем тромбоцитов $< 8,8$ мкм ³ ; лейкоциты $< 5,2 \times 10^9$ /л	13	14,9	23,1
Средний объем тромбоцитов $\geq 8,8$ мкм ³ ; возраст $< 20,0$ года	12	13,8	8,3
Средний объем тромбоцитов $< 8,8$ мкм ³ ; лейкоциты $\geq 5,2 \times 10^9$ /л	31	35,6	3,2

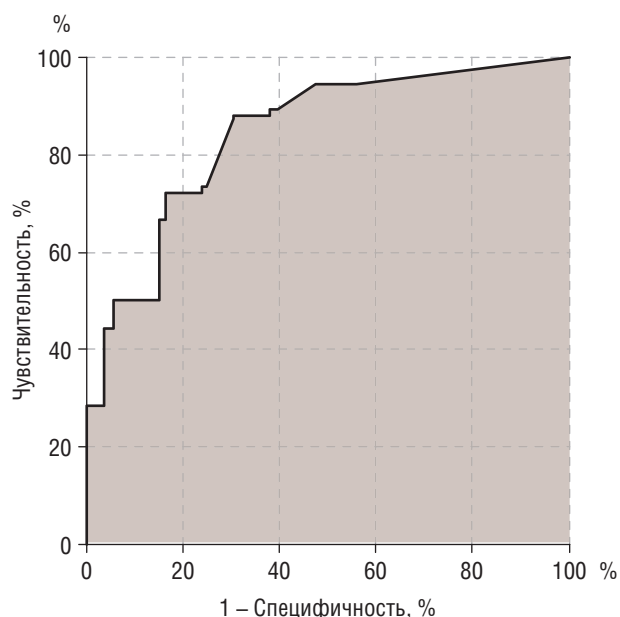


Рис. 2. ROC-кривая для длительности заболевания

Таблица 4. Прогнозные показатели качества построенного «дерева решений» для целевого показателя «длительное заболевание»

Точка отсечения	23,1%
AuROC	0,84
Чувствительность	88,2%
Специфичность	69,5%
Эффективность	78,9%

понимание состояния его здоровья, доверие к рекомендациям медицинских работников и жизненные обстоятельства, позволяющие следовать рекомендациям [18].

Согласно результатам проведенного однофакторного анализа рисков увеличения длительности болезни, можно констатировать, что ключевыми статистически значимыми факторами является наличие фарингита и температуры

тела $\geq 37,0$ °C с уровнями абсолютного риска возникновения длительного заболевания $>51,4\%$. Наличие одного из ключевых прогностических факторов свидетельствует, что заболевание будет протекать длительно (более 5 дней) с осложненным течением.

Значение AuROC, равное 0,84, позволяет считать, что смоделированное «дерево решений» имеет среднее прогностическое качество. Если оценка риска $\geq 23,1\%$, то следует считать, что результат может быть положительным (длительность болезни будет менее 5 дней), в противном случае – отрицательным (длительность болезни будет более 5 дней). При таком подходе в 88,2% случаев правильно идентифицирован положительный результат и в 69,5% случаев – отрицательный результат. Диагностическая ценность показателя оценивается как средняя (AUC=0,84); чувствительность и специфичность – высокие (88,2 и 69,5% соответственно).

Такое сочетание чувствительности и специфичности модели логистической регрессии позволяет рекомендовать использование данного метода на ранних этапах диагностического процесса. Рассмотренный в статье метод математико-статистического моделирования прогноза вероятности развития длительного течения ОРВИ, разработанный с использованием логистического регрессионного анализа, а также метод ROC-анализа в роли оценки его эффективности являются лишь малой долей того потенциала, который представляют современные инновационные технологии [19].

Вероятность ложноотрицательных прогнозов по длительности течения болезни в силу высокой чувствительности отобранного предиктора минимальна, что дает возможность сразу после проведения анализа сформировать группы риска по вероятности развития длительного течения респираторного заболевания. Применение на практике разработанного метода диагностики способно оптимизировать оказание медицинской помощи больным с острыми респираторными вирусными заболеваниями длительного течения.

Персонализированные инструменты, соответствующие уникальным характеристикам отдельного пациента, могут быть важным компонентом персонализированного медицинского обслуживания [20].

СВЕДЕНИЯ ОБ АВТОРАХ

Рычкова Ольга Александровна (Olga A. Rychkova) – доктор медицинских наук, заведующий кафедрой инфекционных болезней, аллергологии и иммунологии ФГБОУ ВО «Тюменский ГМУ» Минздрава России, главный внештатный аллерголог-иммунолог Департамента здравоохранения Тюменской области, Тюмень, Российская Федерация
E-mail: richkovaoa@mail.ru

Кашуба Эдуард Алексеевич (Eduard A. Kashuba) – доктор медицинских наук, профессор кафедры инфекционных болезней, аллергологии и иммунологии ФГБОУ ВО «Тюменский ГМУ» Минздрава России, Тюмень, Российская Федерация
E-mail: kashuba.ed@gmail.com

Рукина Наталья Юрьевна (Natalia Yu. Rukina) – кандидат медицинских наук, доцент кафедры биологической химии и клинической лабораторной диагностики ФГБОУ ВО ДВГМУ Минздрава России, Хабаровск, Российская Федерация
E-mail: rukinanat@mail.ru

Сизов Дмитрий Анатольевич (Dmitry A. Sizov) – врач-инфекционист ГБУЗ «Салехардская окружная клиническая больница», Салехард, Российская Федерация
E-mail: king_rhobar_3@inbox.ru

Перминов Николай Владимирович (Nikolai V. Perminov) – врач-терапевт 301-го военного клинического госпиталя ВВО, Хабаровск, Российская Федерация
E-mail: 301vkg-1@mail.ru

ЛИТЕРАТУРА

1. Халафян А.А. STATISTICA 6. Статистический анализ данных. 3-е изд. Москва : Бинум-Пресс, 2007. 512 с.
2. Лабораторная диагностика гриппа и других ОРВИ методом полимеразной цепной реакции : клинические рекомендации. Москва, 2016. 62 с.
3. Morice A., McGarvey L., Pavord I. Recommendations for the management of cough in adults // *Thorax*. 2006. Vol. 61. P. i1-i24.
4. Adams P.F., Hendershot G.E., Marano M.A.; Centers for Disease Control and Prevention/National Center for Health Statistics. Current estimates from the National Health Interview Survey, 1996 // *Vital Health Stat* 10. 1999. Vol. 200. P. 1-203.
5. Azoulay E., Mokart D., Kouatchet A. et al. Acute respiratory failure in immunocompromised adults // *Lancet Respir. Med*. 2019. Vol. 7. P. 173-186. DOI: [https://doi.org/10.1016/S2213-2600\(18\)30345-X](https://doi.org/10.1016/S2213-2600(18)30345-X)
6. To K.K., Hung I.F., Li I.W., Lee K.L., Koo C.K., Yan W.W. et al. Delayed clearance of viral load and marked cytokine activation in severe cases of pandemic H1N1 2009 influenza virus infection // *Clin. Infect. Dis*. 2010. Vol. 50, N 6. P. 850-859.
7. Costa A.L.P., Neto O.A.R., Silva-Júnior A.C.S. Conditioners of the infectious diseases dynamics // *Estação Científica (UNIFAP)*. 2019. Vol. 8. P. 9-23.
8. Garnett G.P., Cousens S., Hallett T.B., Steketee R., Walker N. Mathematical models in the evaluation of health programmes // *Lancet*. 2011. Vol. 378. P. 515-525. DOI: [https://doi.org/10.1016/S0140-6736\(10\)61505-X](https://doi.org/10.1016/S0140-6736(10)61505-X)
9. Kermack W., McKendrick A. A contribution to the mathematical theory of epidemics // *Proc. R. Soc. Lond. Ser. A Contain. Pap. Math. Phys. Character*. 1927. Vol. 115. P. 700-721.
10. Kermack W.O., McKendrick A.G. Contributions to the mathematical theory of epidemics. II. The problem of endemicity // *Proc. R. Soc. Lond. Ser. A Contain. Pap. Math. Phys. Character*. 1927. Vol. 138. P. 55-83.
11. Heesterbeek H., Anderson R.M., Andreasen V., Bansal S., De Angelis D., Dye C. et al. Modeling infectious disease dynamics in the complex landscape of global health // *Science*. 2015. Vol. 347. Article ID aaa4339. DOI: <https://doi.org/10.1126/science.aaa4339>
12. Peng L., Xie P., Tang Z., Liu F. Modeling and analyzing transmission of infectious diseases using generalized stochastic Petri Nets // *Appl. Sci*. 2021. Vol. 11. Article ID 8400. DOI: <https://doi.org/10.3390/app111884>
13. Флетчер Р., Флетчер С., Вагнер Э. Клиническая эпидемиология. Основы доказательной медицины : пер. с англ. Москва : Медиа Сфера, 1998. 352 с.
14. Изакова Н.Б., Капустина Л.М. Применение метода деревьев классификации для сегментирования корпоративных клиентов на промышленном рынке // *Russian Journal of Management*. 2016. Т. 4, вып. 1 (19). С. 47-53.
15. Шитиков В.К., Мاستицкий С.Э. Классификация, регрессия и другие алгоритмы Data Mining с использованием R. Тольятти, 2017. 351 с. Электронная книга. URL: <https://github.com/ranalytics/data-mining>
16. Айвазян С.А., Мхитарян В.С. Теория вероятностей и прикладная методика и методология обработки информации. Москва : ЮНИТИ, 2001.
17. StatSoft, Inc. Электронный учебник по статистике. Москва : StatSoft, 2012. URL: <http://www.statsoft.ru/home/textbook/default.htm>
18. Ziegelstein R.C. Personomics: the missing link in the evolution from precision medicine to personalized medicine // *J. Pers. Med*. 2017. Vol. 7. P. 11.
19. Григорьев С.Г., Лобзин Ю.В., Скрипченко Н.В. Роль и место логистической регрессии и гос-анализа в решении медицинских диагностических задач // *Журнал инфектологии*. 2016. Т. 8, № 4. С. 36-45. DOI: <https://doi.org/10.22625/2072-6732-2016-8-4-36-45>
20. Athar M.W., Record J.D., Martire C., Hellmann D.B., Ziegelstein R.C. The effect of a personalized approach to patient education on heart failure self-management // *J. Pers. Med*. 2018. Vol. 8, N 4. P. 39. DOI: <https://doi.org/10.3390/jpm8040039> PMID: 30486472; PMCID: PMC6313701.

REFERENCES

1. Khalafyan A.A. STATISTICS 6. Statistical analysis of data. 3rd ed. Moscow: Binom-Press, 2007: 512 p. (in Russian)
2. Laboratory diagnosis of influenza and other acute respiratory viral infections by polymerase chain reaction. Clinical guidelines. Moscow, 2016: 62 p. (in Russian)
3. Morice A., McGarvey L., Pavord I. Recommendations for the management of cough in adults. *Thorax*. 2006; 61: i1-24.
4. Adams P.F., Hendershot G.E., Marano M.A.; Centers for Disease Control and Prevention/National Center for Health Statistics. Current estimates from the National Health Interview Survey, 1996. *Vital Health Stat* 10. 1999; 200: 1-203.
5. Azoulay E., Mokart D., Kouatchet A., et al. Acute respiratory failure in immunocompromised adults. *Lancet Respir Med*. 2019; 7: 173-86. DOI: [https://doi.org/10.1016/S2213-2600\(18\)30345-X](https://doi.org/10.1016/S2213-2600(18)30345-X)
6. To K.K., Hung I.F., Li I.W., Lee K.L., Koo C.K., Yan W.W., et al. Delayed clearance of viral load and marked cytokine activation in severe cases of pandemic H1N1 2009 influenza virus infection. *Clin Infect Dis*. 2010; 50 (6): 850-9.
7. Costa A.L.P., Neto O.A.R., Silva-Júnior A.C.S. Conditioners of the infectious diseases dynamics. *Estação Científica (UNIFAP)*. 2019; 8: 9-23.
8. Garnett G.P., Cousens S., Hallett T.B., Steketee R., Walker N. Mathematical models in the evaluation of health programmes. *Lancet*. 2011; 378: 515-25. DOI: [https://doi.org/10.1016/S0140-6736\(10\)61505-X](https://doi.org/10.1016/S0140-6736(10)61505-X)
9. Kermack W., McKendrick A. A contribution to the mathematical theory of epidemics. *Proc R Soc Lond Ser A Contain Pap Math Phys Character*. 1927; 115: 700-21.
10. Kermack W.O., McKendrick A.G. Contributions to the mathematical theory of epidemics. II. The problem of endemicity. *Proc R Soc Lond Ser A Contain Pap Math Phys Character*. 1927, 138: 55-83.
11. Heesterbeek H., Anderson R.M., Andreasen V., Bansal S., De Angelis D., Dye C., et al. Modeling infectious disease dynamics in the complex landscape of global health. *Science*. 2015; 347: aaa4339. DOI: <https://doi.org/10.1126/science.aaa4339>
12. Peng L., Xie P., Tang Z., Liu F. Modeling and analyzing transmission of infectious diseases using generalized stochastic Petri Nets. *Appl Sci*. 2021; 11: 8400. DOI: <https://doi.org/10.3390/app111884>
13. Fletcher R., Fletcher S., Wagner E. Clinical epidemiology. Fundamentals of evidence-based medicine. Transl. from English. Moscow: Media Sphera Publishers, 1998: 352 p. (in Russian)
14. Izakova N.B., Kapustina L.M. Application of the classification tree method for segmenting corporate clients in the industrial market. *Russian Journal of Management*. 2016; 4 [1 (19)]: 47-53. (in Russian)
15. Shitikov V.K., Mastitsky S.E. Classification, Regression and Other Data Mining Algorithms Using R. Tol'yatti, 2017: 351 p. E-book. URL: <https://github.com/ranalytics/data-mining> (in Russian)
16. Ayvazyan S.A., Mkhitarayan V.S. Probability theory and applied technique and methodology of information processing. Moscow: YUNITI, 2001. (in Russian)
17. StatSoft, Inc. Electronic textbook on statistics. Moscow: StatSoft, 2012. URL: <http://www.statsoft.ru/home/textbook/default.htm> (in Russian)
18. Ziegelstein R.C. Personomics: the missing link in the evolution from precision medicine to personalized medicine. *J Pers Med*. 2017; 7: 11.
19. Grigor'ev S.G., Lobzin Yu.V., Skripchenko N.V. The role and place of logistic regression and roc-analysis in solving medical diagnostic problems. *Zhurnal infekologii [Journal of Infectology]*. 2016; 8 (4): 36-45. DOI: <https://doi.org/10.22625/2072-6732-2016-8-4-36-45> (in Russian)
20. Athar M.W., Record J.D., Martire C., Hellmann D.B., Ziegelstein R.C. The effect of a personalized approach to patient education on heart failure self-management. *J Pers Med*. 2018. 8 (4): 39. DOI: <https://doi.org/10.3390/jpm8040039> PMID: 30486472; PMCID: PMC6313701.