

МАТЕМАТИЧЕСКИЕ МЕТОДЫ И ИНФОРМАЦИОННЫЕ СИСТЕМЫ В ХИМИЧЕСКОЙ ТЕХНОЛОГИИ

УДК 004.94

ПРОГНОСТИЧЕСКАЯ МОДЕЛЬ ПРОТИВООПУХОЛЕВОГО ЭФФЕКТА ТАРГЕТНЫХ ПРЕПАРАТОВ ИММУНОТЕРАПИИ

В.В. Бурляев^{1,*}, профессор, А.А. Давыденко¹, аспирант, О.М. Николаева¹, аспирант, Л.И. Руссу², научный сотрудник, И.А. Суетина², ведущий научный сотрудник, М.В. Мезенцева², руководитель лаборатории

¹Кафедра информационных систем в химической технологии, Московский технологический университет (Институт тонких химических технологий), Москва, 119571 Россия

²Лаборатория культур тканей, Федеральный научно-исследовательский центр эпидемиологии и микробиологии имени Н.Ф. Гамалеи Минздрава России, Москва, 123098 Россия

*Автор для переписки, e-mail: burliaevv@yandex.ru

На основе использования методов индуктивного вывода предложена прогностическая модель построения гипотез о взаимосвязи комбинации цитокинов с пролиферативной активностью раковых клеток. Модель учитывает синергическое взаимодействие цитокинов и использует последовательное построение логических формул для отбора групп цитокинов, статистический анализ таблиц сопряженности и логическую интеграцию полученных оценок. Реализация предложенной модели в рамках информационной системы прогнозирования противоопухолевого эффекта таргетных препаратов иммунотерапии позволит существенно ускорить научные исследования в этой области.

Ключевые слова: информационная система, таблица сопряженности, таргетные препараты, иммунотерапия, комбинация цитокинов, раковая клетка.

THE PROGNOSTIC MODEL OF ANTITUMOR EFFECT OF TARGETED DRUGS IN IMMUNOTHERAPY

V.V. Burlyaev¹, A.A. Davydenko¹, O.M. Nikolaeva¹, L.I. Russu², I.A. Suetina², M.V. Mezentseva²

¹Moscow Technological University (Institute of Fine Chemical Technologies), Moscow, 119571 Russia

²Gamaleya Federal Research Center of Epidemiology and Microbiology of the Ministry of Health of the Russian Federation, Moscow, 123098, Russia

*Corresponding author e-mail: burliaevv@yandex.ru

A prognostic model for constructing hypotheses about the relationship of combinations of cytokines with the proliferative activity of cancer cells is proposed. The model is based on the use of inductive inference methods. The methodology takes into account the synergistic interaction of cytokines and uses sequential construction of logical formulas for selecting groups of cytokines, a statistical analysis of contingency tables and logical integration of the obtained estimates. Implementation of the proposed method in the information system of forecasting the effect of targeted anticancer drugs in immunotherapy will greatly accelerate research in this area.

Keywords: information system, contingency table, targeted drugs, immunotherapy, combination of cytokines, cancer cell.

Одним из перспективных направлений лечения онкологических заболеваний является комбинирование традиционных методов с иммунотерапией, направленной на снижение токсичности препаратов химиотерапии. Особое значение иммунотерапия приобретает на поздних стадиях заболевания, в частности, помогая преодолевать резистентность опухолевых клеток к химиотерапии. В тех случаях, когда химиотерапию или лучевую терапию невозможно проводить ввиду тяжелого общего состояния пациента или сопутствующей патологии, самостоятельное применение иммунотерапии позволяет остановить развитие болезни и продлить жизнь, сохранив высокое качество жизни больного [1].

Разработанные в России препараты Рефнот и Ингарон показали эффективность при онкологических заболеваниях различной локации и в настоящее время активно внедряются в клиническую практику [2]. Создание новых препаратов иммунотерапии, в том числе таргетных препаратов, действующих непосредственно на клетку-мишень и минимально воздействующих на здоровые клетки, находится на острие современных исследований.

Прогнозирование противоопухолевого эффекта таргетных препаратов иммунотерапии связано с анализом влияния изменений уровней транскрипции специфических белков-цитокинов на изменение пролиферативной активности раковых клеток [3, 4]. Для прогнозирования целесообразно использование информационной системы, позволяющей анализировать большое количество взаимосвязанных факторов. Данные об уровнях транскрипции цитокинов и биологической активности клеток карциномы легкого были получены в ФГБУ «Федеральный научно-исследовательский центр эпидемиологии и микробиологии имени почетного академика Н.Ф. Гамалеи» Министерства здравоохранения Российской Федерации.

Биологические эксперименты

Для экспериментов по культивированию из Коллекции культур тканей ФГБУ ФНИЦЭМ им. Н.Ф. Гамалеи Минздрава России были выбраны клетки А-549 (карцинома легкого человека). Для культивирования клеток использовали стандартную питательную среду Игла MEM-90% с 10% эмбриональной телячьей сывороткой.

Для изучения цитотоксического действия препаратов на культуру клеток применяли МТТ-тест. Механизм и причины изменения функциональной клеточной активности был определен с помощью изучения продукции цитокинов на уровне их транскрипции *in vitro* в клеточных культурах после инкубации клеток с препаратами в течение 24 и 48 ч. Экспрессия генов

интерлейкинов ИЛ-1 β , ИЛ-2, ИЛ-4, ИЛ-6, ИЛ-8, ИЛ-10, ИЛ-12, ИЛ-17, ИЛ-18, фактора некроза опухолей ФНО- α , интерферонов ИФН- α , ИФН- β , ИФН- γ , ИФН- λ 1, ИФН- λ 2, ИФН- λ 3 оценивалась по активности их мРНК. Определение активности мРНК цитокинов в клетках проводили с использованием методов обратной транскрипции и полимеразной цепной реакции (ОТ-ПЦР) [5].

Изменения уровня транскрипции цитокинов оценивались качественно (или есть, или нет). Изменение пролиферативной активности раковых клеток карциномы легкого также оценивалось качественно (есть ли изменения по сравнению с контрольным образцом).

Методическая часть, результаты и их обсуждение

Отметим, что многочисленными исследованиями доказан синергизм действия цитокинов на клетки [6]. Многие иммунные реакции обусловлены согласованным действием нескольких цитокинов. При этом одни цитокины способны усиливать или ослаблять продукцию других цитокинов. Поэтому при анализе влияния уровней транскрипции цитокинов на пролиферативную активность раковых клеток необходимо учитывать синергическое взаимодействие цитокинов. Таким образом, первым этапом исследований является отбор групп цитокинов, влияние которых должно анализироваться совместно. Ряд таких комбинаций выявлен и изучен при исследовании механизмов взаимодействия цитокинов с клетками, другие могут быть сформированы в качестве гипотез в процессе анализа экспериментальных данных.

Поскольку все исследуемые величины оцениваются качественно, их можно рассматривать как логические переменные, которые могут принимать одно из значений истинности («истина», «ложь»). Для анализа зависимости между такими переменными удобно использовать методы индуктивного вывода. В частности, изменение пролиферации можно представить как логическую функцию, зависящую от переменных, описывающих изменение уровня транскрипции цитокинов. По существу, исходные данные можно рассматривать как таблицу истинности такой функции. Известно, что всякую логическую функцию можно представить с помощью операций классической логики высказываний (конъюнкции, дизъюнкции и отрицания) в виде дизъюнктивной нормальной формы. На рис. 1 представлен фрагмент базы экспериментальных данных и пример построения логической формулы на его основе.

Анализ полученной логической формулы позволяет выявить комбинации цитокинов, наиболее перспективные для дальнейших исследований. Однако следует учитывать, что классическая логика позволяет устанавливать лишь «строгие» причин-

Идентификатор	ИЛ-1	ИЛ-6	ИЛ-18	Изменение биологической активности
1	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
2	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
3	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
5	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

$$F = (\text{ИЛ-1}\&\text{ИЛ-6}\&\text{ИЛ-18}) \vee (\text{ИЛ-6}\&\text{ИЛ-18}) \vee (\text{ИЛ-6}) \vee (\text{ИЛ-1}\&\text{ИЛ-18}) \vee (\text{ИЛ-18})$$

Рис. 1. Пример построения логической формулы по таблице, описывающей изменения уровней транскрипции цитокинов и биологическую активность раковых клеток.

но-следственные связи (событие А всегда влечет за собой событие В), в то время как индуктивный вывод предполагает построение и проверку гипотез о возможных причинно-следственных связях (событие А, как правило, влечет за собой событие В) [7]. Поэтому возможна ситуация, когда логическая функция окажется многозначной (при одних и тех же аргументах значение функции может быть различным). Однако и в этом случае можно формировать термы, представляющие собой наиболее перспективные комбинации цитокинов [8].

Для проверки и количественного анализа взаимосвязей между группами цитокинов и пролиферативной активностью раковых клеток применялся статистический анализ.

В качестве базового объекта статистического анализа использовались таблицы сопряженности [9]. Строки таблицы сопряженности соответствуют комбинации цитокинов (фактору), столбцы – значениям пролиферативной активности (отклику) [10]. Введем следующие обозначения:

n_{ij} – элементы таблицы сопряженности;
 $s_i = \sum_{j=1}^2 n_{ij}$ – общее количество измерений, в кото-

рых фактор принимает i -ое значение;

$c_j = \sum_{i=1}^2 n_{ij}$ – общее количество измерений, в ко-

торых отклик принимает j -ое значение;

$N = \sum_{i=1}^2 \sum_{j=1}^2 n_{ij}$ – общее количество измерений.

Общий вид таблицы сопряженности приведен на рис. 2.

		отклик		
		есть	нет	
фактор	есть	n_{11}	n_{12}	s_1
	нет	n_{21}	n_{22}	s_2
		c_1	c_2	N

Рис. 2. Общий вид таблицы сопряженности.

Прежде всего, необходимо выполнить статистическую проверку наличия зависимости активности от каждого фактора. Для проверки требуется рассчитать несколько различных статистических критериев: критерий Пирсона χ^2 , критерий Пирсона χ^2 с

поправкой Йейтса, расчет отношения правдоподобия Вальда ($\Lambda\chi^2$), точный критерий Фишера.

Если полученное в результате расчетов значение уровня значимости (p) меньше 0.05 при выбранной доверительной вероятности 95%, то фактор влияет на отклик.

Далее рассчитываются статистические коэффициенты, оценивающие силу связи между каждым фактором и активностью: коэффициент Крамера, коэффициент сопряженности Пирсона, коэффициент связи Юла, коэффициент связи по отношениям преобладаний.

Помимо оценки взаимосвязи между фактором и активностью, следует рассчитать два широко распространенных в медицинских исследованиях коэффициента:

- относительный риск $RR = \frac{n_{11} (n_{21} + n_{22})}{n_{21} (n_{11} + n_{12})}$,
- отношение шансов $OR = \frac{n_{11} n_{22}}{n_{21} n_{12}}$.

Для примера рассмотрим результаты статистических расчетов по 78 экспериментам, выполненным в Федеральном научно-исследовательском центре эпидемиологии и микробиологии имени почетного академика Н.Ф. Гамалеи. Для первоначального исследования были выбраны 6 факторов – комбинаций цитокинов (рис. 3) из публикаций, посвященных цитокинотерапии рака легких, безотносительно к раковым клеткам А-549 [11, 12].

Результаты расчета по таблицам сопряженности для фактора 1 представлены на рис. 4.

Результаты расчетов для всех 6 факторов приведены в табл. 1.

Как видно из таблицы, последняя комбинация (фактор 6) оказалась статистически не значимой для пролиферации клеток А-549.

Разумеется, эти результаты должны быть подтверждены и дополнены при поступлении новых экспериментальных данных.

Таблицы сопряженности, примененные для статистического однофакторного анализа, могут быть использованы для учета совместного влияния выбранных факторов. Для подобного многофакторного анализа применим методику неоднородной последова-

Факторы	Логические формулы
1	(ИНФγ&ИЛ-4) & ((-ИЛ-1) ∨ (ИЛ-6))
2	(ИНФγ & ИНФα) & ((-ИЛ-1) ∨ (ИЛ-12))
3	(ИНФλ ∨ ИНФα) & ((-ИЛ-1) & (ИЛ-2))
4	(ИНФγ ∨ ИЛ-12) & ((-ИЛ-10) ∨ (ИЛ-8))
5	((ИЛ-8) ∨ (ИЛ-6)) & ((ИЛ-10) ∨ (ИЛ-4))
6	((ИЛ-4) & (ИЛ-6) & ((ИЛ-2)&ИНФα))

Рис. 3. Комбинации цитокинов для статистических расчетов.

		результат		суммы по строкам
		есть	нет	
фактор	есть	28	14	42
	нет	14	22	36
суммы по столбцам		42	36	78
Вычисляемые статистики				
Критерий X-квадрат		$\chi^2=$	6,019	Значим!
Критерий правдоподобия		$Y^2=$	6,088	Значим!
Критерий X-квадрат с поправкой Йетса		$J\chi^2=$	4,953	Значим!
<i>При p= 0,05 при наличии связи эти критерии должны быть больше 3,841</i>				
Точный критерий Фишера		Fet=	0,009	Значим!
<i>При наличии связи должен быть меньше 0,05</i>				
Коэффициент связи ν Крамера		$\nu=$	0,278	Сила взаимосвязи средняя
Коэффициент связи c Пирсона		$c=$	0,268	Сила взаимосвязи средняя
Коэффициент связи Q Юла		$Q=$	0,517	Ноль при отсутствии связи
		нижняя граница ДИ	0,177	Не может быть отрицательной!
		верхняя граница ДИ	0,857	
Коэффициент связи по отношению преобладаний		$Ln(Psi)=$	1,115	Ноль при отсутствии связи
Относительный риск		RR=	1,714	
Логарифм относительного риска		Ln(RR)=	0,539	Ноль при отсутствии связи!
		нижняя граница логарифма ДИ	0,077	Не может быть отрицательной!
		верхняя граница логарифма ДИ	1,001	
Отношение шансов		OR=	3,143	
Логарифм отношения шансов		Ln(RR)=	1,145	Ноль при отсутствии связи!
		нижняя граница логарифма ДИ	0,217	Не может быть отрицательной!
		верхняя граница логарифма ДИ	2,073	

Рис. 4. Расчет статистических показателей для фактора 1.

Таблица 1. Результаты расчета статистических показателей

Номер фактора	Критерий X-квадрат	Критерий правдоподобия	Критерий Йейтса	Критерий Фишера точный	Коэффициент Крамера	Коэффициент Пирсона	Коэффициент Юла	Коэффициент преобладаний	Относительный риск RR	Отношение шансов OR	Значимость фактора
1	6.01	6.09	4.96	10 ⁻³	0.28	0.27	0.52	1.12	1.72	3.14	значим
2	7.12	6.33	5.81	0.008	0.21	0.20	0.53	3.21	2.25	3.23	значим
3	8.3	8.07	7.25	0.003	0.22	0.27	0.48	2.84	2.21	2.88	значим
4	4.64	4.19	3.84	0.022	0.17	0.16	0.45	2.63	1.97	2.62	значим
5	3.42	3.048	2.39	0.046	0.143	0.141	0.445	2.63	1.94	2.60	значим
6	0.71	0.68	0.38	0.12	0.065	0.065	0.18	1.46	1.31	1.43	не значим

тельной процедуры [13]. Эта методика основана на методе последовательного статистического анализа Вальда и позволяет построить прогностическую шкалу баллов путем вычисления специальных прогностических коэффициентов и оценки их информативности.

Разработка прогностической шкалы по этой методике состоит из нескольких этапов.

Сначала вычисляются прогностические коэффициенты (ПК) для каждого значимого фактора. Прогностические коэффициенты для каждого из значимых факторов рассчитываются по формулам:

для первого столбца таблицы сопряжения (см. рис. 2)

$$ПК(+)=3 \cdot \ln\left(\frac{n_{11}/c_1}{n_{12}/c_2}\right)$$

и для второго столбца таблицы сопряжения

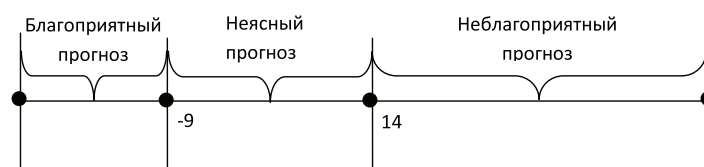


Рис. 5. Прогностическая шкала пролиферативной активности раковых клеток.

Затем при появлении новых экспериментальных данных проверяют в них наличие или отсутствие факторов из табл. 1, после чего вычисляют алгебраическую сумму ПК и выясняют, в какую зону прогноза она попадает.

Ввиду малого объема экспериментальной выборки, оценку прогностических свойств модели было решено провести методом скользящего контроля «выбрасыванием по одному» [14]. Средняя квадратичная ошибка прогноза по этому методу составила 23.6%. Эта ошибка может быть уменьшена при увеличении объема экспериментальных данных и добавлении новых значимых факторов.

Таким образом, с помощью разработанной прогностической шкалы для новых экспериментальных данных можно сделать прогноз пролиферативной активности раковых клеток в зависимости от комбинаций цитокинов.

Основные этапы методики прогнозирования

Список литературы:

1. Козлов В.А., Черных Е.Р. // Бюллетень СО РАМН. 2004. Т. 2. С.113–121.
2. Брюзгин В.В., Платинский Л.В. // Современная онкология. 2014. Т. 16. № 11. С. 2–7.
3. Симбирцев А.С. // Медицинский академический журнал. 2013. Т. 13. № 1. С. 7–22.
4. Лопатина О.А., Мезенцева М.В., Бакланова О.В., Подчерняева Р.Я., Егоров В.В. // Вопросы биол., мед. и фарм. химии. 2014. № 7. С. 22–28.
5. Ершов Ф.И., Мезенцева М.В., Васильев А.Н.,

$$ПК(-)=3 \cdot \ln\left(\frac{n_{21}/c_1}{n_{22}/c_2}\right)$$

Результаты расчетов по этим формулам округляют до целых значений.

Затем выбираются пороговые значения, определяющие три зоны прогностической шкалы: зона неблагоприятного прогноза (где вероятность пролиферативной активности раковых клеток велика), зона с неясным (иногда его называют сомнительным) прогнозом (где такая вероятность равна 50 на 50) и зона благоприятного прогноза (где эта вероятность мала).

Для этого по номограммам [13] были выбраны верхний (14) и нижний (-9) пороги, разбивающие всю прогностическую шкалу на 3 зоны (рис. 5).

пролиферативной активности раковых клеток в зависимости от комбинаций цитокинов, информация, необходимая для выполнения каждого из этапов, и ее результаты схематично представлены на рис. 6.

Выводы

Предложена модель прогнозирования пролиферативной активности раковых клеток на основе гипотез о взаимосвязи комбинации цитокинов. Методика учитывает синергическое взаимодействие цитокинов и использует последовательное построение логических формул для отбора групп цитокинов, статистический анализ таблиц сопряженности и построение прогностической шкалы. Реализация предложенной методики в рамках информационной системы прогнозирования противоопухолевого эффекта таргетных препаратов иммунотерапии позволит существенно ускорить научные исследования в этой области.

References:

1. Kozlov V.A., Chernyh E.R. // Bulletin SB RAMS. 2004. V. 2. P. 113–121. (in Russ.)
2. Bryuzgin V.V., Platinsky L.V. // Modern Oncology. 2014. V.16. № 11. P. 2–7. (in Russ.)
3. Simbirtsev A.S. // Medical Academic Journal. 2013. V. 13. № 1. P. 7–22. (in Russ.)
4. Lopatina O.A., Mezentseva M.V., Baklanova O.V., Podchernyaeva R.J., Egorov V.V. // Questions of biological, medical and pharmaceutical chemistry. 2014. № 7. P. 22–28. (in Russ.)

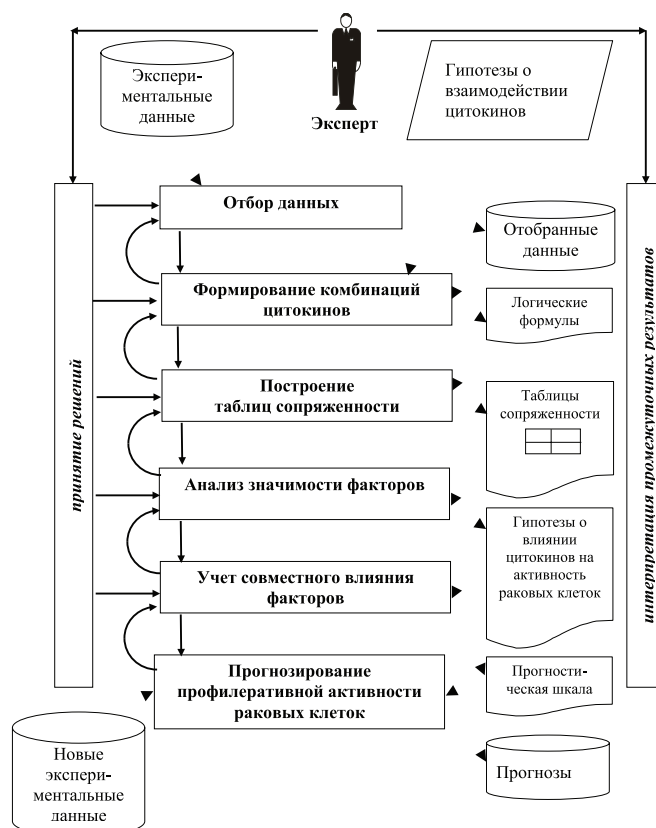


Рис. 6. Методика прогнозирования пролиферативной активности раковых клеток в зависимости от комбинации цитокинов.

Щербенко В.Э., Наровлянский А.Н. // Ведомости Научного центра экспертизы и государственного контроля лекарственных средств. 2002. № 1 (9). С. 26–29.

6. Телетаева Г.М. // Практическая онкология. 2007. Т. 8. № 4. С. 211–218.

7. Вагин В.Н., Головина Е.Ю., Загорянская А.А., Фомина М.В. М.: Физматлит, 2008. 712 с.

8. Голенков В.В., Степанова М.Д., Самодумкин С.А., Гулякина Н.А. Статистические основы индуктивного вывода. Минск: БГУИР, 2009. 202 с.

9. Аптон Г. Анализ таблиц сопряженности. М.: Финансы и статистика, 1982. 143 с.

10. Бурляев В.В., Давыденко А.А., Сутина И.А., Мезенцева М.В. // Интеграл. 2015. № 1-2. С. 26–27.

11. Гельцер Б.И., Маркелова Е.В., Просекова Е.В., Кочеткова Е.А. // Терапевтический архив. 2002. Т. 74. № 11. С. 94–99.

12. Бережная Н.М. // Аллергология и иммунология. 2004. Т. 5. № 3. С. 368–369.

13. Лапач С.Н., Чубенко А.В., Бабич П.Н. Статистические методы в медико-биологических исследованиях с использованием Excel. Киев: МОРИОН, 2000. 320 с.

14. Воронцов К.В. Комбинаторный подход к оценке качества обучаемых алгоритмов // Математические вопросы кибернетики / под ред. О.Б. Лупанова. М.: Физматлит, 2004. Т. 13. С. 5–36.

5. Ershov F.I., Mezentseva M.V., Vasiliev A.N., Shcherbenko V.E., Narovlyansky A.N. // Bulletin of Scientific Center of Expertise and State Control of Medicines. 2002. № 1 (9). P. 26–29. (in Russ.)

6. Teletaeva G.M. // Practical Oncology. 2007. V. 8 № 4. P. 211–218. (in Russ.)

7. Vagin V.N., Golovina E.Yu., Zagoryanskaya A.A., Fomina M.V. Reliable and credible conclusion in intelligent systems. M.: FIZMATLIT, 2008. 712 p. (in Russ.)

8. Golenkov V.V., Stepanova M.D., Samodumkin S.A., Gulyakina N.A. Statistical bases of inductive inference. Minsk: BSUIR, 2009. 202 p. (in Russ.)

9. Upton G. Analysis of contingency tables. M.: Finance and Statistics, 1982. 143 p. (in Russ.)

10. Burlyayev V.V., Davydenko A.A., Suetin I.A., Mezentseva M.V. // Integral. 2015. №1-2. P. 26–27. (in Russ.)

11. Geltser B.I., Markelova E.V., Prosekova E.V., Kochetkova E.A. // Therapeutic Archives. 2002. V. 74. № 11. P. 94–99. (in Russ.)

12. Berezhnaya N.M. // Allergology and Immunology. 2004. V. 5. № 3. P. 368–369. (in Russ.)

13. Lapach S.N., Chubenko A.V., Babich P.N. Statistical methods in biomedical research using Excel. Киев: MORION, 2000. 320 p. (in Russ.)

14. Vorontsov K.V. Combinatorial approach to assessing the quality of training algorithm / Ed. K.V. Vorontsov // Mathematical problems of cybernetics. M.: FIZMATLIT, 2004. V. 13. P. 5–36. (in Russ.)