

Г. С. Лбов<sup>1</sup>, Г. Л. Полякова<sup>1</sup>, В. Н. Бахвалова<sup>2</sup>, О. В. Морозова<sup>3</sup>

<sup>1</sup> Институт математики СО РАН  
просп. Академика Коптюга, 4, Новосибирск, 630090, Россия

<sup>2</sup> Институт систематики и экологии животных СО РАН  
ул. Фрунзе, 11, Новосибирск, 630091, Россия

<sup>3</sup> Институт химической биологии и фундаментальной медицины СО РАН  
просп. Академика Лаврентьева, 8, Новосибирск, 630090, Россия  
E-mail: lbov@math.nsc.ru

## ИССЛЕДОВАНИЕ ВЛИЯНИЯ ПРИРОДНЫХ ФАКТОРОВ НА ЗАБОЛЕВАЕМОСТЬ КЛЕЩЕВЫМ ЭНЦЕФАЛИТОМ \*

Предложен метод анализа эмпирической информации, представленной в виде многомерных временных рядов, отражающих астрофизические, природные характеристики и показатели заболеваемости населения клещевым энцефалитом. Результаты статистического анализа представлены в виде логических закономерностей, отражающих причинно-следственные связи объекта исследования.

*Ключевые слова:* клещевой энцефалит, метод анализа эмпирической информации, логические закономерности.

Заболеваемость людей клещевым энцефалитом (КЭ) подвержена цикличности. Максимальная заболеваемость населения на отдельных, эндемичных по КЭ, территориях обычно отмечается 1–2 года, затем в течение 1–5 лет следует период спада с последующей стабилизацией на минимальном уровне, продолжающейся 6–7 лет, и новым подъемом до максимального уровня [1–4]. Следует отметить, что периодичность характерна не только для общего уровня заболеваемости, но и для соотношения клинических форм [4].

До настоящего времени уровень иммунизации инактивированными вакцинами против КЭ в большинстве регионов России [3] не превышает нескольких процентов от общей численности населения и, следовательно, не оказывает существенного влияния на заболеваемость. Исключения составляют Свердловская область и Республика Алтай, в которых уровни иммунизации населения составляют 65 [5] и 40 % [6] соответственно. При этом именно в этих эндемичных областях уровни заболеваемости

КЭ остаются высокими и превышают ежегодные среднероссийские показатели в 6–10 раз [5; 6].

При обсуждении причин изменчивости заболеваемости людей КЭ большое значение придается природным факторам, в частности температуре воздуха в зимний период, которая определяет численность популяций мелких млекопитающих и соответственно их эктопаразитов, и влажности [7]. Помимо метеорологических факторов, паразитарная система, включающая вирус КЭ, его позвоночных резервуарных хозяев и членистоногих переносчиков, как и все экологические системы в биосфере Земли, находится под влиянием активности Солнца [8]. В цикличности солнечной активности обычно выделяют 11-летний период. Менее известны другие циклы, самый длительный из которых 2 400 лет, а самый короткий – период собственных колебаний Солнца – около 5 мин [9]. Под действием изменяющегося солнечного ультрафиолетового излучения на ионосферу, а солнечного ветра на магнитосферу могут циклически изме-

\* Работа выполнена при частичной финансовой поддержке Междисциплинарного интеграционного проекта СО РАН № 83.

Авторы благодарят за предоставленный статистический материал по астрофизическим и природным факторам директора Сибирского регионального научно-исследовательского гидрометеорологического института д-ра физ.-мат. наук В. Н. Крупчатникова, по заболеваемости КЭ – сотрудников Управления Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека по Новосибирской области.

няться электрическое, переменное электромагнитное поля, инфразвук и даже уровни радиоактивности. К классическим космофизическим индексам традиционно относят числа Вольфа (число пятен и их групп на диске) и индексы магнитной активности. Из новых индексов особо выделяют дни прохождения планеты вблизи границ секторов межпланетного магнитного поля [9]. В странах Западной Европы для прогнозирования заболеваемости людей КЭ используется географическая информационная система (Geographic information systems, GIS), включающая информацию о температуре, влажности, соотношении численности личинок и нимф клещей и других параметрах [10].

**Цель** исследования заключалась в анализе влияния метеорологических и астрофизических факторов на заболеваемость клещевым энцефалитом населения Новосибирской области.

### Материал и методы

Для исследования использовались следующие данные:

- температура воздуха, относительная влажность воздуха и температура точки росы в приземном слое в Новосибирской области;
- данные о солнечной радиации – прямая радиация на горизонтальную и перпендикулярную поверхности, отраженная радиация;
- данные о солнечной активности (число Вольфа);
- показатели заболеваемости клещевым энцефалитом (число заболевших на 100 тыс. населения).

Исследование влияния каждого из указанных природных факторов на заболеваемость КЭ заключалось в статистической обработке 12 временных рядов. Каждый ряд представлял собой набор среднемесячных значений показателей каждого из перечисленных 7 природных факторов и годовых значений заболеваемости людей в течение 18 лет (с 1991 по 2008 г.). Указанная информация получена отдельно для Новосибирской области, г. Новосибирска и Советского района областного центра. Таким образом, каждому природному фактору соответствовала таблица из 12 столбцов

(среднемесячные значения фактора) (переменные  $X_1, \dots, X_{12}$ ) и 18 строк (количество лет периода наблюдений). Для каждой строки указывается количество заболевших за год (переменная  $Y$ ). Требовалось найти статистические закономерности (взаимосвязи) между всеми  $X_1, \dots, X_{12}$  и  $Y$  для каждого фактора. В силу того что природных факторов 7, а рассматриваемых регионов 3, возникла 21 задача выявления закономерностей.

Отметим, что решение подобного рода задач связано с следующими особенностями.

1. Отсутствие априорной информации о распределениях в пространстве переменных (характеристик).

2. В трудноформализуемых областях исследователь вынужден включать большое число потенциально полезных переменных (характеристик) из-за сложности изучаемого явления.

3. Малое число наблюдений (объектов), как правило, сравнимое с числом переменных, хотя из всех известных теоретических исследований следует, что число наблюдений должно превышать в 100–1 000 раз число переменных.

4. Исследователя при изучении сложных объектов нередко интересует не только решение, дающее хорошее качество прогноза, но и сама форма представления такого решения для получения информации о внутренних причинно-следственных связях между характеристиками изучаемых объектов.

При решении статистических задач выявления закономерностей с указанными выше особенностями возникает ряд проблем.

Во-первых, задачи приходится решать в условиях отсутствия априорной информации о виде функций распределения. Любое предположение (например, о нормальности распределения, линейной регрессионной зависимости, независимости переменных, марковости процесса) ставит вопрос о соответствии выбранного предположения истинным ограничениям. Как вводить предположения и какие?

Во-вторых, в условиях малого числа наблюдений и высокой размерности пространства переменных возникает проблема статистической устойчивости получаемых

решений. Из теоретических исследований [11; 12] следует, что чем более сложные функциональные зависимости используются для построения решений, чем больше переменных и чем меньше число наблюдений (объем выборки), тем больше вероятность получения решения, отличающегося от оптимального. Суть проблемы устойчивости статистических решений заключается в следующем. С одной стороны, сильное ограничение на класс решений ставит вопрос об адекватности наших предположений истинному распределению: чем больше такое несоответствие, тем хуже решение. С другой стороны, чем более сложный класс функций используется при малом объеме выборки, тем выше вероятность получить недостоверное решение. Так, например, может оказаться, что линейная функция, заданная на всем множестве переменных, будет менее информативной по сравнению с линейной функцией, заданной на подмножестве этих переменных.

Таким образом, при построении решения необходимо стремиться к максимальной сложности используемого класса решений (для ослабления ограничений на истинное распределение), но при этом сложность класса не должна превышать некоторого порога, задаваемого объемом выборки. При малом объеме выборки класс решений должен иметь малую меру сложности. При увеличении объема выборки этот класс должен позволять постепенно увеличивать свою сложность, вплоть до получения оптимального решения при произвольном распределении. Класс решений, обладающий таким свойством, будем называть универсальным классом. Вопрос о соотношении сложности используемого класса решений и объема выборки – наиболее важный и трудный в общих теоретических исследованиях, связанных с построением решений на основе ограниченной эмпирической информации.

По указанным причинам, а именно при отсутствии информации о виде распределения, наличии малого количества наблюдений (объектов) относительно числа переменных (характеристик), традиционные статистические методы оказываются мало подходящими при решении указанных задач. Подтверждением является результат

исследования влияния вакцинации на заболеваемость КЭ населения Новосибирской области, в котором для выявления зависимости потребовалось предварительное разбиение на временные интервалы, на каждом из которых была обнаружена линейная зависимость.

Как показывают теоретические и экспериментальные исследования [11; 12], построение логико-вероятностных моделей изучаемых явлений на основе такой информации оказывается достаточно перспективным направлением. Модели представляют собой список логических закономерностей, обладающих достаточно высокой прогнозирующей способностью. Кроме того, в отличие от классических методов статистического анализа результаты в рамках указанной модели представляются на языке, близком к естественному языку логических суждений, что облегчает интерпретацию результатов. Анализ информации сводится к выбору нескольких закономерностей из большого количества логических высказываний. Под логическим высказыванием понимается произвольное утверждение на значениях характеристик (переменных) явления (объекта), например, типа «если пятая числовая переменная принимает значение в некотором указанном интервале и одновременно двадцатая переменная (значениями которой являются имена) принимает одно из двух имен, то целевая переменная принимает первое значение (первый класс, образ)». Данное высказывание является логическим высказыванием длины 2. Ясно, что таких высказываний различной длины и использующих различные сочетания исходных переменных может быть достаточно велико (например,  $10^{10}$  и более). Поэтому при малом числе наблюдений (малом объеме выборки) вероятность того, что при одном наблюдении некоторое высказывание будет истинным практически равна нулю. Но если на выборке, например, некоторое высказывание истинно на 10 наблюдениях из 20, то, по-видимому, оно с большей вероятностью является логической закономерностью. Постулируется, что изучаемое явление характеризуется лишь небольшим числом закономерностей (несколько десятков). Для решения поставленной задачи был использован алгоритм, основанный на ме-

тоде [11], который обнаруживает все закономерности с заданной прогнозирующей способностью.

При решении каждой из 21 задачи выявления закономерностей между природными факторами и заболеваемостью клещевым энцефалитом имеем таблицу данных:

$$T = \{x^i, y^i\},$$

где  $x^i = (x_1^i, \dots, x_{12}^i), i = 1, \dots, 18$ .

Так как объем выборки ( $N = 18$ ) является весьма малым, а размерность пространства переменных относительно велика (12), статистически надежные закономерности могут быть получены лишь при огрублении статистического материала. В данном случае диапазон значений каждой переменной разбивается на ряд интервалов. Интервалы переменной  $Y$  назовем образами. Из исходной таблицы  $T$  создаем таблицу  $v = \{z^i, u^i\}$ , где  $z$  и  $u$  – соответствующие интервалы переменных  $x$  и  $y$ . Для каждого сочетания переменных  $z_1, \dots, z_{12}$  перебираются все наборы интервалов рассматриваемого подмножества переменных (логические высказывания, в данном случае конъюнкции). Пример конъюнкции интервалов:  $S = (z_5 = 2) \wedge (z_7 = 1) \wedge (z_{11} = 3)$ . Естественно, ее можно записать через границы интервалов указанных переменных, что и делается при описании результатов изучения влияния природных факторов на заболеваемость клещевым энцефалитом. Логические закономерности определяются отдельно для каждого образа-интервала  $Y$ . Для этого образ с номером  $s, s = 1, \dots, k$ , назовем первым образом, а объединение всех остальных образов – вторым образом.

### Результаты исследования и обсуждение

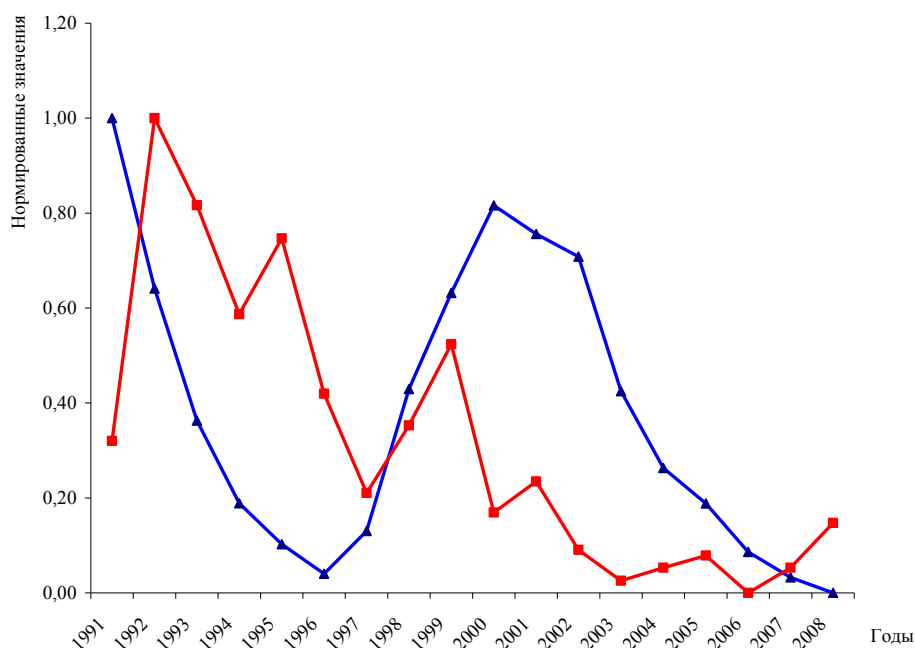
Проведен математический анализ влияния 6 взаимозависимых среднемесячных показателей (температуры воздуха, относительной влажности воздуха и температуры точки росы в приземном слое, а также прямой радиации на горизонтальную поверхность, прямой радиации на перпендикулярную поверхность, отраженной радиации), измеренных на метеостанции вблизи Советского района г. Новосибирска, и среднемесячных, среднегодовых чисел Вольфа на уровне заболеваемости КЭ населения Советского района, г. Новосибирска и Ново-

сибирской области в течение 18-тилетнего периода.

Следовательно, каждому фактору соответствовал 12-мерный временной ряд с 18 отсчетами по числу лет. Для каждого временного отсчета (года) указан показатель заболеваемости клещевым энцефалитом на 100 тыс. населения. Из-за сложности статистической задачи (многомерность, нестационарность, малое число наблюдений), возникла необходимость в «огрублении» статистической информации (использовалось небольшое число интервалов значений переменных, включая и показатель заболеваемости). Значения показателя заболеваемости были разделены на 2 диапазона: 1-й диапазон соответствовал низкому уровню заболеваемости (от 2 до 19 человек на 100 тыс. населения), 2-й диапазон – высокому (от 20 до 62 заболевших на 100 тыс. населения). Это дало возможность обнаружить логические закономерности, как правило, включая конъюнкции значений не более трех переменных для прогноза уровня заболеваемости. В связи с тем что метеорологические и астрофизические факторы измеряли вблизи Советского района (п. Огурцово), а колебания температуры и влажности на территории Новосибирской обл. весьма существенны, наиболее достоверными являются результаты проведенного анализа для Новосибирского научного центра (ННЦ).

В результате анализа данных 18-летнего периода наблюдений с 1991 по 2008 г. выявлены следующие закономерности. Высокий уровень заболеваемости КЭ в Советском районе г. Новосибирска наблюдали при выполнении следующих условий:

- температура воздуха: в апреле от  $-1,5$  до  $3$  °С и в июле от  $19$  до  $22,2$  °С или в июле от  $19$  до  $22,2$  °С, в августе от  $15$  до  $18,7$  °С и в октябре от  $2,1$  до  $5,6$  °С;
- температура точки росы в приземном слое: в марте от  $-17$  до  $-6,2$  °С и в мае от  $-1$  до  $3$  °С или в апреле от  $-7,8$  до  $-2$  °С, в июле от  $12$  до  $16$  °С и в декабре от  $-21$  до  $-10$  °С;
- относительная влажность воздуха: в феврале от  $72,5$  до  $83,5$  % и в декабре от  $79,8$  до  $85,0$  % или в феврале от  $72,5$  до  $83,5$  %, и в апреле от  $58,0$  до  $73,0$  % и в декабре от  $79,8$  до  $85,0$  %;



Динамика солнечной активности и заболеваемости КЭ.  
Синей кривой обозначены числа Вольфа (нормированные),  
красной – заболеваемость КЭ в Советском р-не г. Новосибирска (нормированная)

- прямая солнечная радиация ( $10^6$  Дж/с·м<sup>2</sup>) на горизонтальную поверхность: в мае от 780 до 1 380 S' и в августе от 591 до 1 190 S' или в мае от 780 до 1 380 S', июне от 940 до 1 670 S' и в октябре от 170 до 420 S';

- прямая солнечная радиация ( $10^6$  Дж/с·м<sup>2</sup>) на перпендикулярную поверхность: в марте от 800 до 1 570 S и в июне от 1 525 до 1 900 S или в марте от 800 до 1 570 S, в апреле от 285 до 2 070 S и в июне от 1 525 до 1 900 S;

- отраженная солнечная радиация ( $10^6$  Дж/с·м<sup>2</sup>): в апреле от 246 до 620 Rk, в мае от 287 до 430 Rk и в сентябре от 150 до 250 Rk или в марте от 720 до 1 070 Rk, в мае от 287 до 430 Rk и в августе от 591 до 1 190 Rk;

- числа Вольфа: в мае от 30 до 90 и июле от 30 до 135 или в феврале от 55 до 135, в марте от 40 до 85 и в декабре от 50 до 110.

Известно, что 6 анализируемых природных факторов являются взаимозависимыми и определяемыми активностью Солнца. Для обнаружения причинно-следственных связей между активностью Солнца и уровнем заболеваемости КЭ приведена их динамика с 1991 по 2008 г. (рис.). Для упрощения анализа значения показателей нормированы

от 0 до 1. Нормировка данных (x) выполнялась по формуле:  $(x - \min(x)) / (\max(x) - \min(x))$ .

Приведены совмещенные графики нормированных чисел Вольфа и нормированных значений заболеваемости КЭ (см. рис.). Из анализа приведенных графиков можно сделать следующие выводы.

1. За исключением 1991 г. межгодовые изменения заболеваемости населения Новосибирской области КЭ и солнечной активностью были в основном однонаправленными, но в период с 1992 по 1999 г. изменения (увеличение или уменьшение) значения показателя солнечной активности на один год предшествовали аналогичным изменениям показателей заболеваемости, а в 2000–2008 и, особенно в 2000–2007 гг., колебания значений показателей были практически синхронными.

2. Если активность Солнца, измеряемая в числах Вольфа, изменялась от 2,8 до 33,3, то наблюдали низкий уровень заболеваемости КЭ. Эта закономерность выполнялась для 10 из 18 лет.

При высокой активности Солнца с числами Вольфа в диапазоне от 33,3 до 145,8

регистрировали высокий уровень заболеваемости. Эта закономерность выполнялась для 8 из 18 лет.

### Заключение

Влияние метеорологических и астрофизических факторов на уровень заболеваемости населения КЭ может быть не только прямым, но и опосредованным, поскольку вирус, вызывающий это заболевание, является компонентом паразитарной системы, включающей также его позвоночных и беспозвоночных резервуарных хозяев. Неблагоприятные для иксодовых клещей гидро-термические условия, такие как излишняя сухость или переувлажнение, могут приводить к снижению численности клещей [13]. Быстрое падение температуры осенью предыдущего года может приводить к увеличению численности личинок клещей, что учитывается в географической информационной системе в Западной Европе [10].

Многoletний мониторинг паразитарной системы для клещевого энцефалита в антропогенном природном очаге лесопарковой зоны Советского района г. Новосибирска [1; 14; 15] показал, что, вопреки утвердившемуся в эпидемиологической службе представлению, ни численность, ни вирусофорность клещей не оказывают непосредственного влияния на уровни заболеваемости людей. Напротив, более высокие эти показатели часто соответствуют относительно низкому уровню заболеваемости людей. В последние годы появились сообщения, что отсутствие корреляции между показателями заболеваемости и численности и зараженности клещей вирусом характерно и для других эндемичных территорий [2–4; 6; 7]. Для коротких временных интервалов (3–6 лет) корреляция между заболеваемостью, обилием и вирусофорностью клещей может быть высока, но, по многолетним данным, положительной зависимости не наблюдается [2]. Объяснить отсутствие зависимости можно, скорее всего, тем, что вирусофорность – это лишь показатель частоты зараженности клещей, который не позволяет судить ни о количестве возбудителя в инфицированных особях, ни о его свойствах, включая соотношение различных генетических вариантов вируса в

одном изоляте. В то же время свойства вируса КЭ также претерпевают циклические изменения, коррелирующие с заболеваемостью людей [15].

Несмотря на относительно близкое географическое расположение Новосибирской и Иркутской областей [7], Республики Алтай [6] и приблизительно одинаковую солнечную активность, динамика заболеваемости населения КЭ существенно отличается, что доказывает влияние различных природных факторов на данную паразитарную систему. Ранее описана роль климатических факторов в динамике заболеваемости населения г. Иркутска КЭ [7]. Выраженная положительная корреляция наблюдалась между заболеваемостью людей и температурой воздуха в мае, июле, августе и сентябре, что совпадает с анализом данных по ННЦ в Новосибирской области.

### Список литературы

1. Добротворский А. К., Бахвалова В. Н., Харитонова Н. Н., Сапегина В. Ф. Динамика параметров паразитарной системы клещевого энцефалита в условиях северной лесостепи Приобья // Сиб. экол. журн. 1994. Т. 1, № 4. С. 369–375.
2. Коренберг Э. И., Ковалевский Ю. В. Основные черты экоэпидемиологии клещевого энцефалита // Вестник инфектологии и паразитологии [Электронный ресурс]. Режим доступа: <http://www.nedug.ru/lib/lit/infect/01oct/infect145/infect.htm> (дата обращения 28.01.2010).
3. Злобин В. И. Эпидемиологическая обстановка и проблемы борьбы с клещевым энцефалитом в Российской Федерации // Бюл. сибирской медицины. 2006. № 5, прил. 1. С. 16–23.
4. Команденко Н. И., Жукова Н. Г. Некоторые дискуссионные вопросы проблемы клещевого энцефалита // Бюл. сибирской медицины. 2006. № 5, прил. 1. С. 57–62.
5. Прохорова О. Г., Романенко В. В., Злобин В. И. Сравнительная характеристика иммунологической активности вакцин клещевого энцефалита, используемых в ходе кампании массовой вакцинации населения Свердловской области // Эпидемиология и вакцинопрофилактика. 2006. Т. 29, № 4. С. 33–36.

6. Щучинова Л. Д. Эпидемиологический надзор и контроль инфекций, передающихся клещами, в Республике Алтай: Автореф. дис. ... канд. мед. наук. Омск, 2009.

7. Коротков Ю. С., Никитин А. Н., Антонова А. М. Роль климатических факторов в многолетней динамике заболеваемости населения г. Иркутска клещевым энцефалитом // Бюл. ВСНЦ СО РАМН. 2007. Т. 55, № 3. С. 126–130.

8. Владимировский Б. М., Темурьянц Н. А. Влияние солнечной активности на биосферу. М., 2000.

9. Владимировский Б. М. Солнечная активность и жизнь // Химия и жизнь. 2009. № 1. С. 10–13.

10. Mantke O. D., Schadler R., Niedrig M. A survey on cases of tick-borne encephalitis in European countries // Eur. Surveill. 2008. Vol. 13. P. 17–27.

11. Лбов Г. С. Методы обработки разнотипных экспериментальных данных. Новосибирск, 1981.

12. Лбов Г. С., Старцева Н. Г. Логические решающие функции и вопросы статистической устойчивости решений. Новосибирск, 1999.

13. Коротков Ю. С., Кисленко Г. С., Буренкова Л. А., Рудникова Н. А., Карань Л. С.

Пространственная и временная изменчивость зараженных клещей *Ixodes ricinus* и *Ixodes persulcatus* возбудителем болезни Лайма в Московской области // Паразитология. 2008. № 6. С. 441–451.

14. Бахвалова В. Н., Добротворский А. К., Панов В. В., Морозова О. В., Матвеева В. А., Мак В. В., Мошкин М. П., Евсиков В. И. Временная изменчивость компонентов паразитарной системы клещевого энцефалита и их связь с заболеваемостью людей // Материалы Всерос. науч. конф. «Современные научные и прикладные аспекты клещевого энцефалита» (к 70-летию открытия вируса клещевого энцефалита). М., 2007. С. 140–141.

15. Бахвалова В. Н., Морозова О. В., Морозов И. В. Свойства популяции вируса клещевого энцефалита, циркулировавшего в 1980–2006 гг. на территории Новосибирской области // Бюл. СО РАМН. 2007. № 4. С. 41–48.

Материал поступил в редколлегию 25.01.2010

**G. S. Lbov, G. L. Polyakova, V. N. Bakhvalova, O. V. Morozova**

#### **The Role of The Climatic Factors in Tick-Born Encephalitis Disease's Indices**

In this paper the method of analysis of empirical information represented in the form of multidimensional time series is suggested. These time series reflect astrophysical, climatic features and tick-born encephalitis disease's indices. Results of statistical analysis are represented in the form of logical rules reflected cause-effect relations of object under investigation.

*Keywords:* tick-born encephalitis, method of analysis of empirical information, logical rules.