

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ЭПИДЕМИЧЕСКИХ ПРОЦЕССОВ И ОЦЕНКА ИХ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК

*Б.М. Десятков, А.И. Бородулин, С.С. Котлярова, Н.А. Лантева,
М.Ю. Марченко, А.Н. Шабанов*

ФГУН ГНЦ ВБ «Вектор» Роспотребнадзора, пос. Кольцово Новосибирской области
Центр Госсанэпиднадзора Новосибирской области, Новосибирск

Проведено сравнение рассчитанных по математической модели и зарегистрированных значений инфицированных индивидуумов во время эпидемий гриппа и кори. Не во всех случаях наблюдается удовлетворительное совпадение исследуемых характеристик. Выполнены оценки дисперсии числа инфицируемых. Расчеты показали, что дисперсия может достигать значительных величин и ее необходимо учитывать при анализе эпидемических процессов.

Ключевые слова: моделирование эпидемического процесса, дисперсия, эпидемии гриппа и кори.

The comparison of the values calculated according to the mathematical model and those recorded for infected individuals during influenza and measles epidemics was performed. Satisfactory coincidence of the studied characteristics is not observed in all the cases. The dispersion of the number of infected individuals was evaluated. The calculations showed that the dispersion can reach significant values and should be taken into account in the analysis of epidemic processes.

Keywords: modeling of epidemic process, dispersion, influenza and measles epidemics.

Введение

В основе современных методов описания динамики развития эпидемических процессов, положены подходы, разработанные в 60–70 годах прошлого века О.В. Барояном и Л.А. Рвачевым [1–3]. Используемые в этих детерминистических моделях уравнения описывают процесс развития эпидемии без учета статистической природы эпидемических процессов. Они дают нам оценку среднего ожидаемого количества заболевших индивидуумов. В то же время на эти характеристики оказывает влияние множество случайных известных и неизвестных факторов, приводящих к тому, что фактические данные о протекшей эпидемии являются случайной величиной. Вследствие этого при практическом использовании прогнозов необходимо иметь представление о возможном статистическом разбросе оценок числа заболевших – дисперсии, которая характеризует отклонение фактического числа заболевших индивидуумов от среднего значения. Неучет этого явления может привести к неправильным, грубым оценкам ре-

сурсов, необходимых для проведения противо-эпидемических и лечебно-профилактических мероприятий.

Большинство подходов к статистическому описанию эпидемических процессов сталкиваются со значительными техническими трудностями, связанными с незамкнутостью уравнений [4]. В соответствии с этим современные статистические модели эпидемических процессов лишь в той или иной степени приближаются к последовательному описанию данных сложных процессов (см., например, [5, 6]). В работах [7, 8] предложен новый метод решения данной проблемы и получены замкнутые уравнения для определения статистических характеристик эпидемического процесса.

В данной работе проводится анализ этого метода и вычисляются первые и вторые моменты числа чувствительных и инфицируемых индивидуумов с использованием информации о сезонных вспышках гриппа в г. Новосибирске 2001–2005 годов [9] и данных о динамике эпидемий гриппа [10] и кори [11].

Материалы и методы

В соответствии с [7, 8] система уравнений для моделирования эпидемического процесса и

определения вторых моментов (дисперсии числа чувствительных и инфицированных) записывается следующим образом:

$$\frac{d\langle S \rangle}{dt} = -v[\langle S \rangle \langle I \rangle + \langle \hat{S} \hat{I} \rangle] ; \quad (1)$$

$$\frac{d\langle I \rangle}{dt} = v[\langle S \rangle \langle I \rangle + \langle \hat{S} \hat{I} \rangle] - w\langle I \rangle .$$

$$\begin{aligned} \frac{d\sigma_S^2}{dt} &= -2v[\langle S \rangle \langle \hat{S} \hat{I} \rangle + \sigma_S^2 \langle I \rangle] ; \\ \frac{d\sigma_I^2}{dt} &= 2v[\langle S \rangle \sigma_I^2 + \langle \hat{S} \hat{I} \rangle \langle I \rangle] - w\sigma_I^2 ; \end{aligned} \quad (2)$$

$$\frac{d\langle \hat{S} \hat{I} \rangle}{dt} = -v[\langle S \rangle \sigma_I^2 + \langle \hat{S} \hat{I} \rangle \langle I \rangle - \langle S \rangle \langle \hat{S} \hat{I} \rangle - \sigma_S^2 \langle I \rangle] - w\langle \hat{S} \hat{I} \rangle .$$

Здесь $\langle S \rangle$ и $\langle I \rangle$ – математическое ожидание числа чувствительных и инфицированных индивидуумов, σ_S^2 и σ_I^2 – дисперсии чувствительных и инфицированных, v – частота заражения индивидуума, отнесенная к одному индивидууму из группы населения $\langle I \rangle$ и отражающая весь комплекс наблюдаемых и ненаблюдаемых явлений эпидемического процесса – скученность населения, интенсивность перемещения, патоген-

ность возбудителей, климатические факторы и др.; w – частота выбывания индивидуума из группы $\langle I \rangle$. Угловыми скобками обозначена процедура усреднения по статистическому ансамблю, «крышечками» сверху обозначены пульсации – отклонение случайной величины от математического ожидания.

Для решения системы уравнений (1), (2) необходимо задать начальные условия:

$$\begin{aligned} \langle S(t=t_0) \rangle &= S_0; \langle I(t=t_0) \rangle = I_0; \sigma_S^2(t=t_0) = \sigma_{S0}^2; \\ \sigma_{I0}^2(t=t_0) &= \sigma_{I0}^2 \text{ и } \langle \hat{S} \hat{I} \rangle \Big|_{t=t_0} = (\rho_{SI} \sigma_S \sigma_I) \Big|_{t=t_0} . \end{aligned}$$

Из (2) видно, что вторые моменты целиком определяются набором первых моментов числа чувствительных и инфицированных, а также коэффициентами v и w .

Систему дифференциальных уравнений (1,2) аппроксимируем системой разностных уравнений

$$\begin{aligned} \langle S(t+\Delta t) \rangle &= -\Delta t v[\langle S(t) \rangle \langle I(t) \rangle + \langle \hat{S} \hat{I}(t) \rangle] ; \\ \langle I(t+\Delta t) \rangle &= \Delta t v[\langle S(t) \rangle \langle I(t) \rangle + \langle \hat{S} \hat{I}(t) \rangle] - \Delta t w \langle I(t) \rangle . \\ \sigma_S^2(t+\Delta t) &= \sigma_S^2(t) - 2\Delta t v[\langle S(t) \rangle \langle \hat{S} \hat{I}(t) \rangle + \sigma_S^2(t) \langle I(t) \rangle] ; \\ \sigma_I^2(t+\Delta t) &= \sigma_I^2(t) + 2\Delta t v[\langle S(t) \rangle \sigma_I^2(t) + \langle \hat{S} \hat{I}(t) \rangle \langle I(t) \rangle] - \Delta t w \sigma_I^2(t) ; \\ \langle \hat{S} \hat{I}(t+\Delta t) \rangle &= \langle \hat{S} \hat{I}(t) \rangle - \Delta t v[\langle S(t) \rangle \sigma_I^2(t) + \\ &+ \langle \hat{S} \hat{I}(t) \rangle \langle I(t) \rangle - \langle S(t) \rangle \langle \hat{S} \hat{I}(t) \rangle - \sigma_S^2(t) \langle I(t) \rangle] - \Delta t w \langle \hat{S} \hat{I}(t) \rangle , \end{aligned} \quad (3)$$

где Δt – шаг по времени, используемый при моделировании эпидемического процесса.

Результаты и обсуждение

Возможности исследуемого метода проверялись на двух типах эпидемических процессов: эпидемия гриппа и эпидемия кори. Было проведено две серии расчетов. В первой серии вычислялись только математические ожидания характеристик эпидемического процесса без учета их дисперсии. Цель этих расчетов – определить прогностические возможности исходной достаточно простой модели. Во второй серии расчетов вычислялись математические ожидания и дисперсии с учетом их взаимного влияния.

При анализе теоретических результатов и сравнении их с реальными данными необходимо учесть следующее. В математической модели (1,2) есть четыре параметра (ν , w , S_0 , I_0), которые, как показали наши многочисленные расчеты, сильно влияют на результаты моделирования. Причем, если значения параметра w ($1/w$ – продолжительность пребывания в группе инфицированных) для каждого заболевания приблизительно известны, то значение параметра ν заранее неизвестно и для различных эпидемий может меняться на порядки. Значение параметра I_0 (число инфицированных в момент начала эпидемии) также приблизительно известно. А вот значение параметра S_0 (начальное число чувствительных) в некоторых случаях неопределенно.

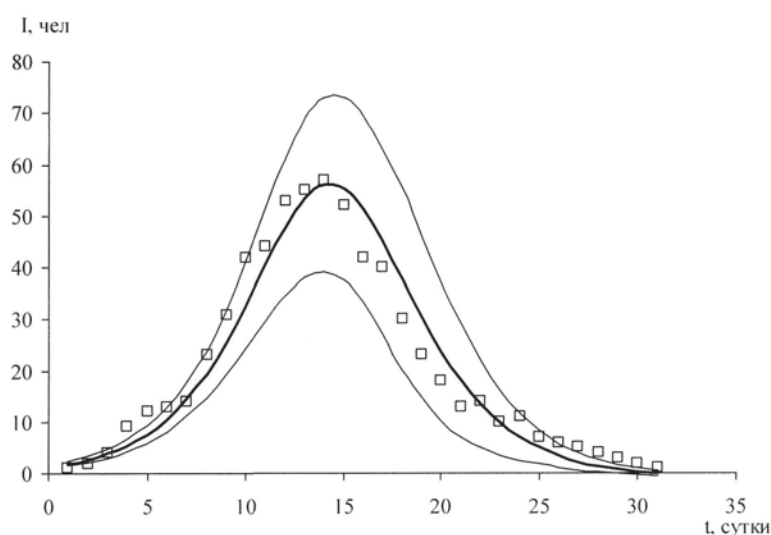


Рис. 1. Рассчитанное (толстая сплошная кривая) и зарегистрированное [10] количество инфицированных гриппом. $S_0 = 360$, $I_0 = 2$, $\nu = 0,24 \cdot 10^{-2}$ (чел. день) $^{-1}$, $w = 0,45$ (день) $^{-1}$. Тонкие линии – число инфицированных плюс (минус) рассчитанное стандартное отклонение

На рис. 1 представлены вычисленные (сплошная толстая линия) и наблюдаемые [10] значения числа инфицированных для эпидемии гриппа в изолированной группе детей (школа-интернат). В этом примере значения $I_0 = 1$ чел. и $S_0 = 365$ чел. достоверны. Хорошее совпадение теоретических и наблюдаемых значений получается при $\nu = 0,0024$ (чел. день) $^{-1}$ и $w = 0,45$ (день) $^{-1}$. Значения этих параметров согласуются с приведенными в [10]. В следующем примере (см. рис. 2) моделируется эпидемия кори в Нью-Йорке в 1962 году [11]. Население Нью-Йорка в том году было 7781984 человека. Однако, как сказано в [11], после тщательного анализа стало ясно, что только 505829 человек не имеют иммунитета от кори, а остальные ранее переболели этой болезнью. Именно это значение $S_0 = 505830$ чел. и было взято нами (так же, как и авторами работы [11]). Как видно из рис. 2, получается удовлетворительное совпадение рассчитанных и наблюдаемых значений динамики числа инфицированных при значениях параметров $I_0 = 21$ чел., $\nu = 0,40 \cdot 10^{-6}$ (чел. день) $^{-1}$ и $w = 0,17$ (день) $^{-1}$, совпадающих с приведенными в [11]. В примерах моделирования эпидемий гриппа, которые прошли в г. Новосибирске в 2001, 2003, 2004 и 2005 годах [9] совпадение рассчитанных и наблюдаемых значений числа инфицированных не столь хорошее. После многочисленных расчетов с различными значениями

упомянутых выше параметров модели наилучшие результаты, представленные на рис. 3–6, были получены при $S_0 = 50000$ чел. для эпидемий 2001, 2003 годов и $S_0 = 10000$ чел. для эпидемий 2004 и 2005 годов и разумных значениях параметров ν и w . Хотя формально мы должны были задавать $S_0 = 1400000$ чел. (население г. Новосибирска). Но для больших значений параметра $S_0 = 500000$ чел., 1000000 чел., 1400000 чел. рассчитанное значение числа инфицированных отличается от наблюдаемых в десятки и сотни раз.

Из приведенных на рис. 3–6 данных видно, что кривые наблюдаемой заболеваемости для разных лет значительно различаются. Ве-

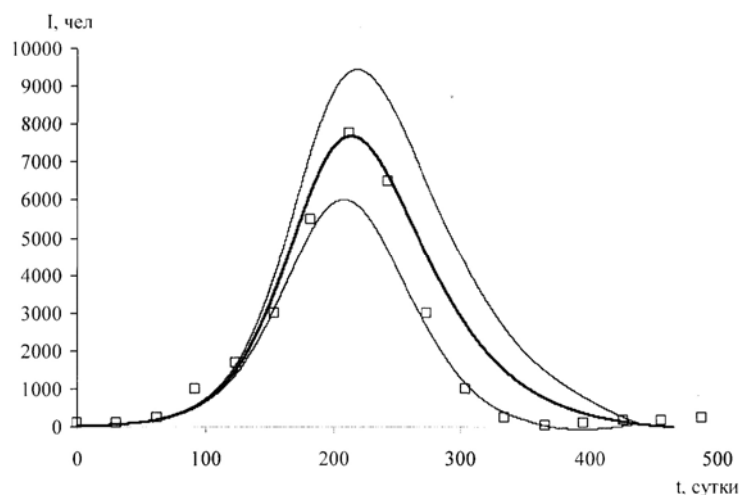


Рис. 2. Рассчитанное (толстая сплошная кривая) и зарегистрированное [2] количество инфицированных корью. $S_0 = 505830$, $I_0 = 21$, $\nu = 0,40 \cdot 10^{-6}$ (чел. день) $^{-1}$ и $w = 0,17$ (день) $^{-1}$. Тонкие линии – число инфицированных плюс (минус) рассчитанное стандартное отклонение

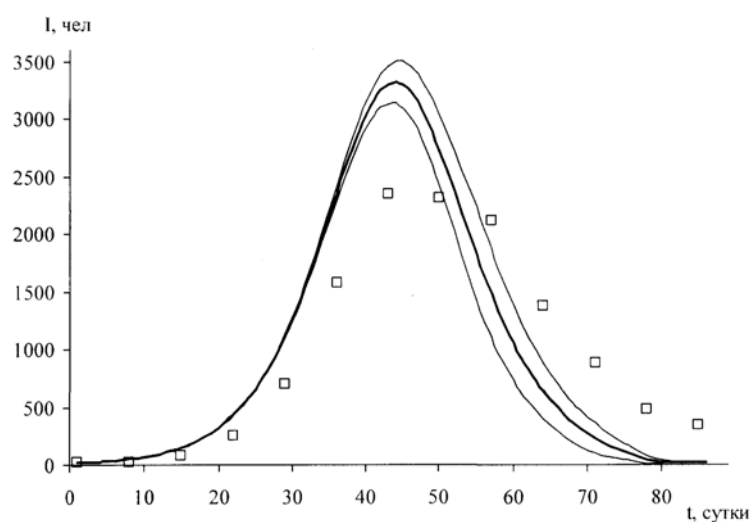


Рис. 3. Рассчитанное (толстая сплошная кривая) и зарегистрированное количество инфицированных во время эпидемии гриппа в Новосибирске в 2001 году. $S_0 = 50000$, $I_0 = 18$, $\nu = 0,10 \cdot 10^{-4}$ (чел. день) $^{-1}$ и $w = 0,33$ (день) $^{-1}$. Тонкие линии – число инфицированных плюс (минус) рассчитанное стандартное отклонение

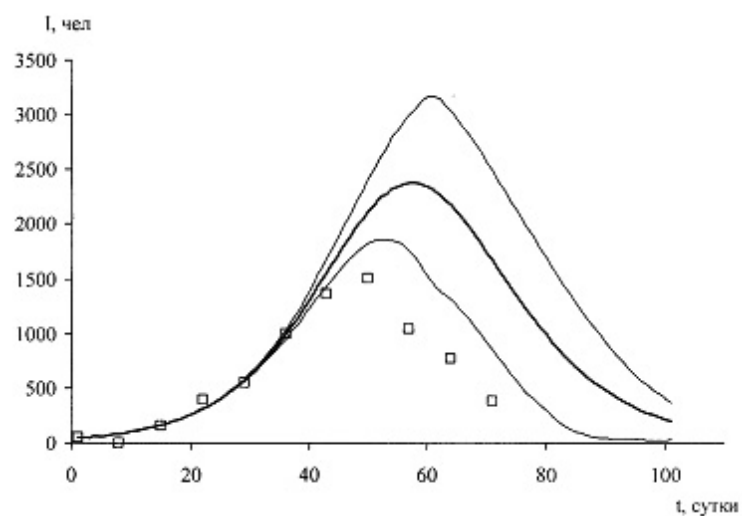


Рис. 4. Рассчитанное (толстая сплошная кривая) и зарегистрированное количество инфицированных во время эпидемии гриппа в Новосибирске в 2003 году. $S_0 = 50000$, $I_0 = 44$, $\nu = 0,70 \cdot 10^{-5}$ (чел. день) $^{-1}$ и $w = 0,25$ (день) $^{-1}$. Тонкие линии – число инфицированных плюс (минус) рассчитанное стандартное отклонение

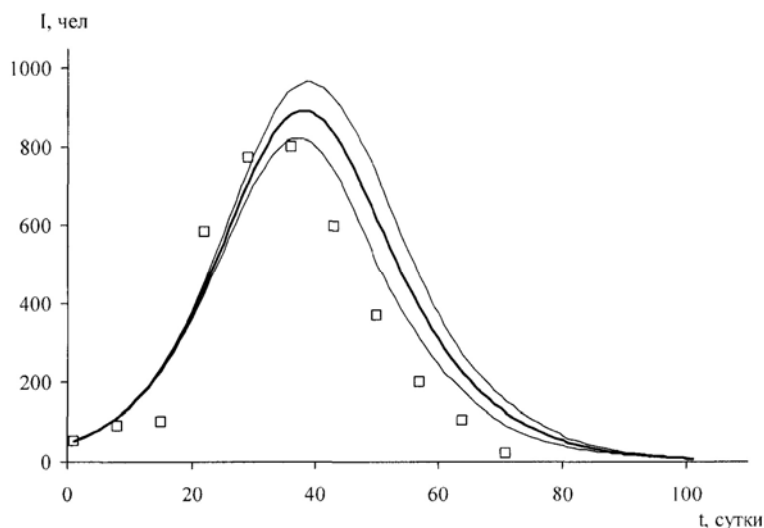


Рис. 5. Рассчитанное (толстая сплошная кривая) и зарегистрированное количество инфицированных во время эпидемии гриппа в Новосибирске в 2004 году. $S_0 = 10000$, $I_0 = 51$, $\nu = 0,32 \cdot 10^{-4}$ (чел. день) $^{-1}$ и $w = 0,20$ (день) $^{-1}$. Тонкие линии – число инфицированных плюс (минус) рассчитанное стандартное отклонение

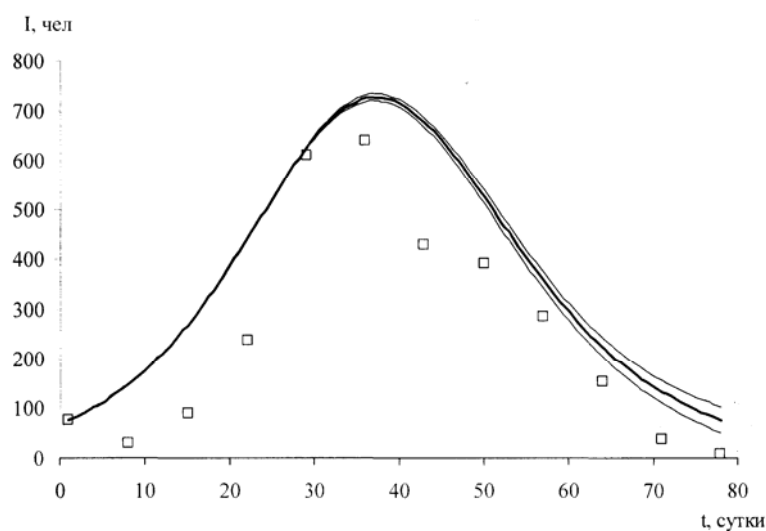


Рис. 6. Рассчитанное (толстая сплошная кривая) и зарегистрированное количество инфицированных во время эпидемии гриппа в Новосибирске в 2005 году. $S_0 = 10000$, $I_0 = 77$, $\nu = 0,30 \cdot 10^{-4}$ (чел. день) $^{-1}$ и $w = 0,20$ (день) $^{-1}$. Тонкие линии – число инфицированных плюс (минус) рассчитанное стандартное отклонение

роятно, эти различия обусловлены разными типами (субтипами) циркулировавшей вирусной инфекции, разным количеством индивидуумов, получивших профилактические прививки, и другими факторами. Очевидно, что в этих примерах практически невозможно заранее определить параметр S_0 – число потенциально чувствительных, которые теоретически могут вступить в контакт с инфицированными, а также параметр ν , отражающий весь комплекс наблюдаемых и ненаблюдаемых явлений эпидемического процесса – скученность населения, интенсивность перемещения, патогенность возбудителей, климатические факторы и др.

С целью анализа поведения параметра ν в эпидемическом процессе была решена обратная задача – по известным из наблюдений значени-

ям числа инфицированных из уравнений (3) для каждого момента времени вычислялся параметр ν . Среднее за все время эпидемии значение этого параметра для эпидемии 2004 года и для $S_0 = 10000$ чел., 100000 чел. и 1000000 чел. равно, соответственно, $0,31 \cdot 10^{-4}$ (чел. день) $^{-1}$, $0,26 \cdot 10^{-5}$ (чел. день) $^{-1}$ и $0,25 \cdot 10^{-6}$ (чел. день) $^{-1}$, а коэффициент вариации этого параметра равен 0,68, 0,90 и 0,91, соответственно. Очевидно, что временная изменчивость этого параметра довольно существенна. Аналогичные результаты получаются и для остальных эпидемий 2001, 2003 и 2005 годов. Принципиально иное поведение этого параметра наблюдается в эпидемиях [10, 11]. Для данных [10] среднее значение и стандартное отклонение его равны $0,27 \cdot 10^{-2}$ (чел. день) $^{-1}$ и 0,48, а для данных [11] эти значе-

Вычисленные значения коэффициента вариации при максимальном значении числа инфицированных

[3]	[4]	2001		2003		2004		2005	
k	k	S ₀ , чел.	k	S ₀ , чел.	k	S ₀ , чел.	k	S ₀ , чел.	k
0,30	0,22	5·10 ⁴	0,07	5·10 ⁴	0,26	1·10 ⁴	0,08	1·10 ⁴	0,03

ния равны $0,47 \cdot 10^{-6}$ (чел. день)⁻¹ и 0,17. Необходимо заметить, что во всех моделях типа [1–3] предполагается, что параметр ν не зависит от времени. Видно, что это предположение удовлетворительно выполняется для данных [10, 11] и заметно хуже выполняется для данных [9].

Из этих примеров следует вывод о необходимости более тщательного анализа числа зарегистрированных инфицированных и разработки специальной процедуры задания параметров ν и S_0 .

Во второй серии расчетов моделировались рассмотренные выше эпидемические процессы с учетом взаимного влияния дисперсии и математического ожидания (3).

На рис. 1–6 пунктирными линиями представлены графики $\langle I \rangle \pm \sigma_I$, а в таблице приведены значения коэффициента вариации $k = \sigma_I / \langle I \rangle$, соответствующие максимальным значениям числа инфицированных индивидуумов. Видно, что в некоторых случаях коэффициент вариации может достигать 30%, а прогнозируемое максимальное значение числа инфицированных с вероятностью 0,68 может находиться в интервале $(0,7 \div 1,3) \langle I \rangle$.

Выводы

1. Вычисленные значения дисперсии числа инфицируемых при эпидемических процессах могут достигать значимых величин и их необходимо учитывать при анализе масштабов эпидемии.

2. Необходимо дальнейшее совершенствование математической модели, в частности, необходим более надежный и объективный метод задания параметров модели.

Список литературы

1. Бароян, О.В., Рвачев, Л.А., Иванников, Ю.Г. Моделирование и прогнозирование эпидемий гриппа на территории СССР. – М., 1977. – 546 с.
2. Бароян, О.В., Рвачев, Л.А. Математика и эпидемиология. – М.: Знание, 1977. – 64 с.
3. Бароян, О.В., Рвачев, Л.А. Прогнозирование эпидемий гриппа в условиях СССР // *Вопр. вирусол.* – 1978. – № 2. – С. 131–137.
4. Бейли, Н. Математика в биологии и медицине. – М.: Мир, 1970. – 326 с.
5. Mode, C.J., Sleeman, C.K. A new design of stochastic partnership models for epidemics of sexually transmitted diseases with stages // *Math. Biosci.* – 1999. – Vol. 156. – P. 95–122.
6. Allen, L.J.S., Kirupaharan, N. Asymptotic dynamics of deterministic and stochastic epidemic models with multiple pathogens // *Int. J. Numeric Anal. and Model.* – 2005. – Vol. 2, No. 3. – P. 329–344.
7. Бородулин, А.И., Десятков, Б.М., Ярыгин, А.А. Статистическая модель эпидемического процесса // *Сиб. ж. индустр. мат.* – 2007. – Т. X, № 2 (30). – С. 23–30.
8. Бородулин, А.И., Десятков, Б.М., Шабанов, А.Н., Ярыгин, А.А. Определение первых и вторых моментов в модели эпидемического процесса Барояна – Рвачева // *Сиб. ж. индустр. мат.* – 2007. – Т. X, № 3 (31). – С. 13–17.
9. Бородулин, А.И., Десятков, Б.М., Шабанов, А.Н. и др. Статистический разброс числа инфицируемых индивидуумов при сезонных вспышках гриппа в г. Новосибирске 2001–2005 годов // В кн.: *Проблемы охраны здоровья населения и обеспечения гигиенической и эпидемиологической безопасности окружающей среды.* – Новосибирск, 2007. – С. 386–390.
10. Stilianakis, N.I., Perelson, A.S., Hayden, F.G. Emergence of Drug Resistance during an Influenza Epidemic: Insights from Mathematical Model // *J. Infect. Diseases.* – 1998. – No. 177. – P. 863–873.
11. Brauer, F., van den Driessche, P., Wu J. *Mathematical Epidemiology.* – Springer-Verlag Berlin Heidelberg, 2008. – 412 p.