pp. 1–19 (2012)

Discrete stochastic model of HIV infection spread within a heterogeneous population

N. V. PERTSEV* and V. N. LEONENKO*

Abstract — A discrete stochastic model of HIV infection spread within a heterogeneous population is presented. A system of high-dimensional stochastic difference equations is used in the model. A system of upper estimates for mathematical expectations of the sizes of groups of individuals in the population is constructed. Sufficient conditions for the extinction of the infection process are obtained in terms of M-matrices. Based on the Monte Carlo method, a numerical algorithm and a simulation code are developed, which allows one to study probabilistic characteristics of the observed variables of the model. Results of numerical experiments are presented.

A pandemia of HIV infection has been progressing since the beginning of 1980th and is one of the most baneful epidemics in the human history. In 2008–2010, about 2.7 millions of new HIV infection cases were diagnosed in the world. In 2010, the total number of registered HIV infection cases was 34 million; 1.8 million of people died from causes related to HIV infection [6]. In order to decrease losses caused by HIV infection, it is necessary to control the prevalence of this infection and to take measures aimed at the detection and treatment of this disease. One of the modern approaches to solution of these problems is connected with application of the method of simulation modelling. For example, a model of HIV infection dynamics in Zambia was developed in [8] with the use of detailed consideration of the structure of sexual contacts of individuals; a model of HIV infection dynamics was constructed in [16] on the base of dynamic networks of individual contacts; a model of HIV infection spread within the homosexual population of San Francisco was described in [17].

Systems of stochastic difference equations with integer-valued variables and their individual-oriented modifications can be used for the development of models simulating the dynamics of HIV infection and other diseases. Such models are a powerful tool for the study of disease spread, they correctly take into account the randomness factor, and peculiarities of particular individuals; these models can be easily adjusted to real data, do not require much time for numerical experiments, and can be studied analytically (see, e.g., [1, 7, 11, 13, 14]). Note also that the ap-

*Omsk Branch of the Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Omsk 644099, Russia

The work was supported by the Program of Fundamental Research of the RAS ‘Modern problems of theoretical mathematics’ and by the Interdisciplinary Integration Project of Siberian Branch of RAS No. 47.

plication of simulation models based on stochastic difference equations allows us to overcome some restrictions existing in the technique of ordinary differential equations. Such restrictions include interpretation complexity of solutions related to the continuity of variables, especially in the case of simultaneous simulation of small and large groups of individuals; impossibility to take into account the discreteness of time caused by the daily (periodical) living rhythm of individuals; neglection of stochastic nature of contacts causing infection of susceptible individuals; the use of only one (exponential) family of distributions describing the total duration of the disease and the duration of its particular stages for a particular individual.

In this paper we consider a system of high-dimensional stochastic difference equations describing the dynamics of HIV infection spread within a heterogeneous population. The stochastic model is based on the results of [12, 15], where the population heterogeneity of Russian regions was considered from the viewpoint of social dysadaptation of individuals. The aim of the development and study of the stochastic model are: (1) description of the population heterogeneity using integer-valued random vectors of arbitrary dimension; (2) formulation of sufficient conditions for the extinction of the infection process; (3) application of numerical experiments to the study of mathematical expectations of the sizes of considered population groups depending on variations of the parameters of the model.

1. Equations of the model

Let us represent the population of a certain region as the groups of individuals $A_1, A_2, \ldots, A_n$, $B_1, B_2, \ldots, B_n$. We assume that the groups $A_1, A_2, \ldots, A_n$ are formed from HIV-susceptible individuals differing in their level of social dysadaptation, for example, socially adapted individuals, individuals with a high risk of development of various pathologies, individuals with a fixed alcoholism or drug dependence. The groups $B_1, B_2, \ldots, B_n$ describe HIV-infected individuals subject to the levels of social dysadaptation indicated above. We assume that individuals are indistinguishable within each group listed above. By $x_i(t), y_i(t)$ we denote the numbers of the individuals in the groups A_i, B_i at time moments $t = 0, 1, 2, \ldots$, $i = 1, 2, \ldots, n$.

We also assume that the sizes of the groups described above vary in time as the result of the following processes:

- transition of individuals from one group to another because of a change in their level of social adaptation, an infection, development of disease, detection and treatment of cases;
- supplement of groups because of immigration and inflow of individuals reaching a particular age;
- decrease caused by natural mortality of individuals, death by the disease, or emigration to other regions.

As an example, Figure 1 presents a scheme of the HIV infection spread model proposed in [12, 15] ($n = 4$).

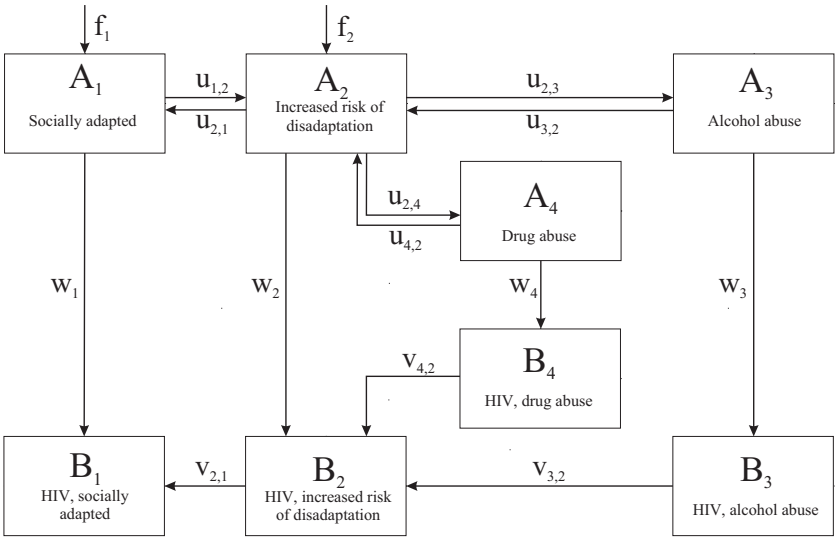

Figure 1. An example of HIV infection spread for $n = 4$ [12, 15].

In order to construct the model, we use a system of stochastic difference equations with integer-valued variables and discrete time. The unit of time is chosen relative to the duration of typical processes occurring to an individual within the model (from one day to one year). We suppose that in each time interval $(t-1;t]$ all individuals behave independently of each other and their behaviour does not depend on events preceding the time moment $t-1$. We fix a time moment $t-1$ and the quantities $x_i(t-1), y_i(t-1)$ of the groups A_i, B_i. The numbers of individuals living at the time moment t are denoted by $\widehat{x}_i(t)$, $\widehat{y}_i(t)$, respectively. We assume that for fixed $i = 1,2,\dots,n$ the random variables $\widehat{x}_i(t)$, $\widehat{y}_i(t)$ have the binomial distribution

$$\widehat{x}_i(t) \sim \mathbf{B}(x_i(t-1),\rho_i), \qquad \widehat{y}_i(t) \sim \mathbf{B}(y_i(t-1),\gamma_i) \tag{1.1}$$

where $\rho_i, \gamma_i \in (0;1)$ are the probabilities of survival from the time moment $t-1$ till the moment t for the individuals from the groups A_i and B_i.

Consider the case when the HIV infection is absent in the population. The system of model equations has the following form:

$$\begin{aligned}
x_1(t) &= \widehat{x}_1(t) - \sum_{k=2}^{n} u_{1,k}(t) + \sum_{l=2}^{n} u_{l,1}(t) + f_1(t) \\
&\cdots\cdots\cdots \\
x_i(t) &= \widehat{x}_i(t) - \sum_{k=1,k\neq i}^{n} u_{i,k}(t) + \sum_{l=1,l\neq i}^{n} u_{l,i}(t) + f_i(t) \\
&\cdots\cdots\cdots \\
x_n(t) &= \widehat{x}_n(t) - \sum_{k=1}^{n-1} u_{n,k}(t) + \sum_{l=1}^{n-1} u_{l,n}(t) + f_n(t) \\
x_i(0) &= x_i^{(0)}, \quad i = 1,2,\dots,n, \quad t = 1,2,\dots\ .
\end{aligned} \tag{1.2}$$

System (1.2) uses the following notations $(i, j = 1, 2, \ldots, n)$:

- $u_{i,j}(t)$ is a nonnegative integer-valued random variable representing for a fixed t the number of individuals of the group A_i passing to the group A_j, $j \neq i$, in the time period $(t-1;t]$;
- $f_i(t)$ is a nonnegative integer-valued random variable representing for a fixed t the inflow of individuals into the group A_i in the time period $(t-1;t]$;
- $x_i^{(0)}$ is a nonnegative integer-valued random variable representing the initial size of the group A_i.

Let us describe the distribution laws of these variables. By $p_{i,j} \in [0;1)$ we denote the probability of the transition of a single individual from the group A_i to the group A_j in the time interval $(t-1;t]$, $i, j = 1, 2, \ldots, n$, $j \neq i$. We assume that for each $i = 1, 2, \ldots, n$ the inequality $\sum_{j=1, j\neq i}^{n} p_{i,j} < 1$, holds. For fixed $\widehat{x}_i(t)$, the random vector

$$u^{(i)}(t) = (u_{i,1}(t), \ldots, u_{i,i-1}(t), u_{i,i}(t), u_{i,i+1}(t), \ldots, u_{i,n}(t)) \tag{1.3}$$

has the multinomial distribution

$$u^{(i)}(t) \sim \mathbf{M}(\widehat{x}_i(t), p_{i,1}, \ldots, p_{i,i-1}, p_{i,i}, p_{i,i+1}, \ldots, p_{i,n})$$

where

$$u_{i,i}(t) = \widehat{x}_i(t) - \sum_{k=1, k\neq i}^{n} u_{i,k}(t), \quad p_{i,i} = 1 - \sum_{j=1, j\neq i}^{n} p_{i,j}. \tag{1.4}$$

Note that the random variable $u_{i,i}(t)$ represents the number of individuals of the groups A_i not passing into any group A_j, $j \neq i$, in the period $(t-1;t]$, i.e., remaining in the group A_i.

For fixed $i = 1, 2, \ldots, n$, the inflow $f_i(t)$ is a random process with given probabilistic characteristics. The variables $x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}$ form the components of a nonnegative integer-valued random vector with a particular distribution law.

Proceed to the system of model equations taking into account the presence of HIV-infected individuals. Fix the time moment $t-1$ and the sizes $\widehat{y}_i(t)$ of living individuals of the groups B_i assuming that $\widehat{y}_1(t) + \widehat{y}_2(t) + \cdots + \widehat{y}_n(t) \neq 0$. Assume also that any individual remaining in the group A_i can be HIV-infected in the period $(t-1;t]$ as the result of a contact with at least one individual from the groups $B_1, B_2, \ldots, B_n$. Each HIV-infected individual from the group A_i supplements the group B_i increasing its size by one, $i = 1, 2, \ldots, n$. The probability of infection for a particular individual from the group A_i is described by a generalized variant of the formula from the Reed–Frost chain-binomial model [5]. Let $r_{i,j} \in [0;1)$ be the probability of a contact of an individual from the group A_i with an individual from

the group B_j, $i,j = 1,2,\ldots,n$, in the time interval $(t-1;t]$ causing a subsequent infection. Then the probability of infection $\mu_i(t)$ of the individual from the group A_i in the period $(t-1;t]$ is given by the formula

$$\mu_i(t) = 1 - \prod_{j=1}^{n} (1 - r_{i,j})^{\widehat{y}_j(t)}, \quad i = 1,2,\ldots,n$$

which can be rewritten in the form

$$\mu_i(t) = 1 - \exp\Big(- \sum_{j=1}^{n} \widehat{y}_j(t) \ln \frac{1}{1 - r_{i,j}} \Big), \quad i = 1,2,\ldots,n. \tag{1.5}$$

By $w_i(t)$ we denote the number of individuals from the group A_i, $i = 1,2,\ldots,n$, infected in the time period $(t-1;t]$. For fixed t, $u_{i,i}(t)$, $\mu_i(t)$, the random variable $w_i(t)$ is described by the binomial distribution

$$w_i(t) \sim \mathbf{B}(u_{i,i}(t), \mu_i(t)), \quad i = 1,2,\ldots,n. \tag{1.6}$$

The number of individuals passing from a fixed group B_i into groups B_j in the time period $(t-1;t]$ is described by the nonnegative integer-valued random variables $v_{i,j}(t)$, $j \neq i$. The probability of an individual from the group B_i passing into the group B_j in the time period $(t-1;t]$ is denoted by $q_{i,j} \in [0;1)$, $i,j = 1,2,\ldots,n$, $j \neq i$. We assume that for each $i = 1,2,\ldots,n$ we have $\sum_{j=1,\, j\neq i}^{n} q_{i,j} < 1$. Denote

$$v_{i,i}(t) = \widehat{y}_i(t) - \sum_{k=1,k\neq i}^{n} v_{i,k}(t), \quad q_{i,i} = 1 - \sum_{j=1,j\neq i}^{n} q_{i,j}.$$

For fixed $\widehat{y}_i(t)$ the random vector

$$v^{(i)}(t) = (v_{i,1}(t),\ldots,v_{i,i-1}(t), v_{i,i}(t), v_{i,i+1}(t),\ldots,v_{i,n}(t))$$

has the multinomial distribution

$$v^{(i)}(t) \sim \mathbf{M}(\widehat{y}_i(t), q_{i,1},\ldots,q_{i,i-1}, q_{i,i}, q_{i,i+1},\ldots,q_{i,n}).$$

We assume that the groups B_i, $i = 1,2,\ldots,n$, have no external inflows. By $y_1^{(0)}, y_2^{(0)},\ldots,y_n^{(0)}$ we denote the components of a nonnegative integer-valued random vector with a given distribution law describing the initial sizes of the groups $B_1, B_2,\ldots,B_n$.

Under the assumptions and notations introduced above, the system of model equations has the form

$$\begin{aligned}
x_i(t) &= \widehat{x}_i(t) - \sum_{k=1,k\neq i}^{n} u_{i,k}(t) - w_i(t) + \sum_{l=1,l\neq i}^{n} u_{l,i}(t) + f_i(t) \\
y_i(t) &= \widehat{y}_i(t) - \sum_{k=1,k\neq i}^{n} v_{i,k}(t) + w_i(t) + \sum_{l=1,l\neq i}^{n} v_{l,i}(t) \\
x_i(0) &= x_i^{(0)}, \quad y_i(0) = y_i^{(0)}, \quad i = 1,2,\ldots,n, \quad t = 1,2,\ldots\ .
\end{aligned} \tag{1.7}$$

System of equations (1.7) determines a Markov random process

$$\mathbf{Z}(t) = (x_1(t), x_2(t), \ldots, x_n(t), y_1(t), y_2(t), \ldots, y_n(t)), \quad t = 1, 2, \ldots$$

with integer-valued nonnegative components. The non-negativity of $x_i(t), y_i(t)$ for all $t = 1, 2, \ldots$, $i = 1, 2, \ldots, n$ follows from the structure of model equations (1.7) and the distribution laws of random variables appearing in this system. A detailed study of the probabilistic characteristics of the process $\mathbf{Z}(t)$ is practically impossible. An exception is the process $\mathbf{Z}(t) = \mathbf{Z}_0(t) = (x_1(t), x_2(t), \ldots, x_n(t), 0, 0, \ldots, 0)$, $t = 1, 2, \ldots$, corresponding to the absence of HIV infection in the population. The process $\mathbf{Z}_0(t)$ is determined by the initial conditions $y_i^{(0)} = 0$, $i = 1, 2, \ldots, n$ (with probability 1). The dynamics of $x_i(t)$, $i = 1, 2, \ldots, n$ may be sufficiently complex in this case, but it admits a study with the use of the standard methods for a Markov chain represented in the form of model (1.2). Another important case is related to the determination of the conditions (restrictions on the model parameters and initial data) providing a complete eradication of HIV infection in the population in the case when the sizes of the groups $B_1, B_2, \ldots, B_n$ are nonzero, i.e., $y_1^{(0)} + y_2^{(0)} + \cdots + y_n^{(0)} > 0$ with probability 1. In other words, here we have to get the conditions providing $\mathbf{P}\{y_i(t) = 0\} \to 1$ for $t \to \infty$ for all $i = 1, 2, \ldots, n$. The following two sections are focused on the solution of this problem.

2. Equations and upper estimates for mathematical expectations of the sizes of groups of individuals

Let the following assumptions hold true:

(H1) The initial sizes of the groups A_i, B_i have finite mathematical expectations $m_i^{(0)} = \mathbf{E}x_i^{(0)}$, $n_i^{(0)} = \mathbf{E}y_i^{(0)}$, $i = 1, 2, \ldots, n$;

(H2) For each $t = 1, 2, \ldots$ the inflows $f_i(t)$ have finite mathematical expectations $\mathbf{E}f_i(t) = f_i^*(t)$, $i = 1, 2, \ldots, n$.

By $m_i(t) = \mathbf{E}x_i(t)$, $n_i(t) = \mathbf{E}y_i(t)$ we denote the mathematical expectations of the sizes of the groups A_i, B_i at the time moments $t = 1, 2, \ldots$, $i = 1, 2, \ldots, n$. First we consider a particular case of the model system of equations (1.2). Fix t, i and calculate $\mathbf{E}x_i(t)$ assuming that there exist finite $\mathbf{E}x_1(t-1), \mathbf{E}x_2(t-1), \ldots, \mathbf{E}x_n(t-1)$. Using distribution law (1.1) and the formula for a conditional mathematical expectation, we get

$$\mathbf{E}\widehat{x}_i(t) = \mathbf{E}\mathbf{E}\{\widehat{x}_i(t) | x_i(t-1)\} = \mathbf{E}\{\rho_i x_i(t-1)\} = \rho_i m_i(t-1).$$

Considering (1.2)–(1.4), we obtain

$$\begin{aligned}\mathbf{E}\{\sum_{k=1,k\neq i}^{n} u_{i,k}(t)\} &= \sum_{k=1,k\neq i}^{n} \mathbf{E}u_{i,k}(t) = \sum_{k=1,k\neq i}^{n} \mathbf{E}(p_{i,k}\rho_i x_i(t-1)) \\ &= \sum_{k=1,k\neq i}^{n} p_{i,k}\rho_i m_i(t-1) = (1-p_{i,i})\rho_i m_i(t-1) \\ \mathbf{E}\{\sum_{l=1,l\neq i}^{n} u_{l,i}(t)\} &= \sum_{l=1,l\neq i}^{n} \mathbf{E}u_{l,i}(t)) \\ &= \sum_{l=1,l\neq i}^{n} \mathbf{E}(p_{l,i}\rho_l x_l(t-1)) = \sum_{l=1,l\neq i}^{n} p_{l,i}\rho_l m_l(t-1).\end{aligned}$$

These relations imply

$$\begin{aligned}m_i(t) &= \rho_i m_i(t-1) - (1-p_{i,i})\rho_i m_i(t-1) + \sum_{l=1,l\neq i}^{n} p_{l,i}\rho_l m_l(t-1) + f_i^*(t) \\ &= \sum_{k=1}^{n} p_{k,i}\rho_k m_k(t-1) + f_i^*(t).\end{aligned}$$

Taking the initial data into account, we get that $m_i(t) = \mathbf{E}x_i(t)$ exist and are finite within model (1.2) and also satisfy the system of linear difference equations

$$\begin{aligned}m_i(t) &= \sum_{k=1}^{n} p_{k,i}\rho_k m_k(t-1) + f_i^*(t) \\ m_i(0) &= m_i^{(0)}, \quad i = 1,2,\ldots,n, \quad t = 1,2,\ldots\ .\end{aligned}$$

Now proceed to model (1.7). Fix t, i and calculate $\mathbf{E}x_i(t)$, $\mathbf{E}y_i(t)$ assuming that

$$\mathbf{E}x_1(t-1), \mathbf{E}x_2(t-1), \ldots, \mathbf{E}x_n(t-1), \mathbf{E}y_1(t-1), \mathbf{E}y_2(t-1), \ldots, \mathbf{E}y_n(t-1)$$

exist and are finite. Based on distribution law (1.6) and on the formula for a conditional mathematical expectation, we get

$$\mathbf{E}w_i(t) = \mathbf{E}\mathbf{E}\{w_i(t)|\mu_i(t))\} = \mathbf{E}\{p_{i,i}\rho_i m_i(t-1)\mu_i(t)\} = p_{i,i}\rho_i m_i(t-1)\mathbf{E}\mu_i(t).$$

It is impossible to write down the expression for $\mathbf{E}\mu_i(t)$ explicitly. Using formula (1.5) and applying Jensen's inequality for mathematical expectations of convex functions of random variables [4], we get the estimates

$$0 \leqslant \mathbf{E}\mu_i(t) \leqslant 1 - \exp\Big(-\sum_{j=1}^{n} \gamma_j n_j(t-1) \ln\frac{1}{1-r_{i,j}}\Big) \leqslant 1$$

which implies the existence and finiteness of $\mathbf{E}\mu_i(t)$. Repeating the calculations presented above, we can easily write

$$m_i(t) = \sum_{k=1}^{n} p_{k,i}\rho_k m_k(t-1) - \mathbf{E}w_i(t) + f_i^*(t)$$
$$n_i(t) = \sum_{j=1}^{n} q_{j,i}\gamma_j n_j(t-1) + \mathbf{E}w_i(t).$$

As the result, we get that the mathematical expectations $m_i(t) = \mathbf{E}x_i(t)$, $n_i(t) = \mathbf{E}y_i(t)$ exist, are finite, and satisfy the following system of recurrence relations:

$$\begin{aligned} m_i(t) &= \sum_{k=1}^{n} p_{k,i}\rho_k m_k(t-1) - p_{i,i}\rho_i m_i(t-1)\mathbf{E}\mu_i(t) + f_i^*(t) \\ n_i(t) &= \sum_{j=1}^{n} q_{j,i}\gamma_j n_j(t-1) + p_{i,i}\rho_i m_i(t-1)\mathbf{E}\mu_i(t) \qquad (2.1) \\ m_i(0) &= m_i^{(0)}, \quad n_i(0) = n_i^{(0)}, \quad i = 1,2,\dots,n, \quad t = 1,2,\dots\ . \end{aligned}$$

Note additionally that all $m_i(t) \geqslant 0$, $n_i(t) \geqslant 0$, $t = 1,2,\dots$. This property follows from the nonnegativity (with probability 1) of the model variables $x_i(t)$, $y_i(t)$, or this can be directly obtained from system (2.1) subject to the conditions $m_i^{(0)} \geqslant 0$, $n_i^{(0)} \geqslant 0$, $f_i^*(t) \geqslant 0$ and the restrictions on the constants that are the parameters of the model.

System (2.1) does not allow us to study the behaviour of $m_i(t), n_i(t)$ directly, because it contains the summands $\mathbf{E}\mu_i(t)$ dependent on $n_1(t-1), n_2(t-1), \dots, n_n(t-1)$ implicitly. Construct estimates for $m_i(t), n_i(t)$ using the system of inequalities

$$\begin{aligned} m_i(t) &\leqslant \sum_{k=1}^{n} p_{k,i}\rho_k m_k(t-1) + f_i^*(t) \\ n_i(t) &\leqslant \sum_{j=1}^{n} q_{j,i}\gamma_j n_j(t-1) + p_{i,i}\rho_i m_i(t-1)\left(1 - \exp\Big(-\sum_{j=1}^{n}\gamma_j n_j(t-1)\ln\frac{1}{1-r_{i,j}}\Big)\right) \qquad (2.2) \\ &\leqslant \sum_{j=1}^{n} q_{j,i}\gamma_j n_j(t-1) + p_{i,i}\rho_i m_i(t-1)\sum_{j=1}^{n}\gamma_j n_j(t-1)\ln\frac{1}{1-r_{i,j}} \end{aligned}$$

and the additional system of difference equations obtained from (2.2) by replacement of all inequalities by equalities. In what follows, assume

$$g_i(z_1,z_2,\dots,z_n) = \sum_{j=1}^{n}\gamma_j z_j \ln\frac{1}{1-r_{i,j}}, \quad z_j \geqslant 0, \quad i,j = 1,2,\dots,n. \qquad (2.3)$$

Proposition 2.1. Let assumptions H1 and H2 hold for model (1.7). Then the mathematical expectations $m_i(t), n_i(t)$ satisfy the estimates

$$0 \leqslant m_i(t) \leqslant \widetilde{m}_i(t), \quad 0 \leqslant n_i(t) \leqslant \widetilde{n}_i(t), \quad i = 1,2,\dots,n, \quad t = 1,2,\dots \tag{2.4}$$

where the variables $\widetilde{m}_i(t), \widetilde{n}_i(t)$ are the solutions to the system of difference equations

$$\begin{gathered}
\widetilde{m}_i(t) = \sum_{k=1}^{n} p_{k,i}\,\rho_k \widetilde{m}_k(t-1) + f_i^*(t) \\
\widetilde{n}_i(t) = \sum_{j=1}^{n} q_{j,i}\,\gamma_j \widetilde{n}_j(t-1) + p_{i,i}\,\rho_i \widetilde{m}_i(t-1)\, g_i(\widetilde{n}_1(t-1), \widetilde{n}_2(t-1), \dots, \widetilde{n}_n(t-1)) \\
\widetilde{m}_i(0) = m_i^{(0)}, \quad \widetilde{n}_i(0) = n_i^{(0)}, \quad i = 1,2,\dots,n, \quad t = 1,2,\dots\ .
\end{gathered} \tag{2.5}$$

Proof. Fix $i = 1,2,\dots,n$. The nonnegativity of $m_i(t), n_i(t)$ has been proved above for all $t = 1,2,\dots$. Proceed to the proof of upper estimates for these variables. By the hypothesis, we have $m_k(0) = \widetilde{m}_k(0), n_k(0) = \widetilde{n}_k(0)$, $k = 1,2,\dots,n$. This gives

$$\begin{aligned}
0 \leqslant \mathbf{E}\mu_i(1) \leqslant & g_i(n_1(0), n_2(0), \dots, n_n(0)) = g_i(\widetilde{n}_1(0), \widetilde{n}_2(0), \dots, \widetilde{n}_n(0)) \\
\widetilde{m}_i(1) - m_i(1) = & \sum_{k=1}^{n} p_{k,i}\,\rho_k\,(\widetilde{m}_k(0) - m_k(0)) \\
& + p_{i,i}\,\rho_i\, m_i(0)\,\mathbf{E}\mu_i(1) = p_{i,i}\,\rho_i\, m_i(0)\,\mathbf{E}\mu_i(1) \geqslant 0 \\
\widetilde{n}_i(1) - n_i(1) = & \sum_{j=1}^{n} q_{j,i}\,\gamma_j\,(\widetilde{n}_j(0) - n_j(0)) \\
& + p_{i,i}\,\rho_i\,\widetilde{m}_i(0)\, g_i(\widetilde{n}_1(0), \widetilde{n}_2(0), \dots, \widetilde{n}_n(0)) - p_{i,i}\,\rho_i\, m_i(0)\,\mathbf{E}\mu_i(1) \\
= & p_{i,i}\,\rho_i\, m_i(0)\,(g_i(n_1(0), n_2(0), \dots, n_n(0)) - \mathbf{E}\mu_i(1)) \geqslant 0.
\end{aligned}$$

Therefore, inequalities (2.4) are valid for $t = 1$. Suppose these inequalities are valid for $t = \tau$ and verify them for $t = \tau + 1$. Fix $i = 1,2,\dots,n$. We have

$$\begin{aligned}
0 \leqslant \mathbf{E}\mu_i(\tau+1) & \leqslant g_i(n_1(\tau), n_2(\tau), \dots, n_n(\tau)) \leqslant g_i(\widetilde{n}_1(\tau), \widetilde{n}_2(\tau), \dots, \widetilde{n}_n(\tau)) \\
\widetilde{m}_i(\tau+1) - m_i(\tau+1) & = \sum_{k=1}^{n} p_{k,i}\,\rho_k\,(\widetilde{m}_k(\tau) - m_k(\tau)) + p_{i,i}\,\rho_i\, m_i(\tau)\,\mathbf{E}\mu_i(\tau+1) \\
& \geqslant p_{i,i}\,\rho_i\, m_i(\tau)\,\mathbf{E}\mu_i(\tau+1) \geqslant 0 \\
\widetilde{n}_i(\tau+1) - n_i(\tau+1) & = \sum_{j=1}^{n} q_{j,i}\,\gamma_j\,(\widetilde{n}_j(\tau) - n_j(\tau)) \\
& + p_{i,i}\,\rho_i\,\widetilde{m}_i(\tau)\, g_i(\widetilde{n}_1(\tau), \widetilde{n}_2(\tau), \dots, \widetilde{n}_n(\tau)) - p_{i,i}\,\rho_i\, m_i(\tau)\,\mathbf{E}\mu_i(\tau+1) \\
& \geqslant p_{i,i}\,\rho_i\, m_i(\tau)\,(g_i(\widetilde{n}_1(\tau), \widetilde{n}_2(\tau), \dots, \widetilde{n}_n(\tau)) - \mathbf{E}\mu_i(\tau+1)) \geqslant 0
\end{aligned}$$

which completes the proof.

3. Sufficient conditions for extinction of infection process

In order to determine conditions providing a complete eradication of HIV infection in a population, we use the Chebyshev inequality [4]: if η is a nonnegative random variable with the mathematical expectation $0 < \mathbf{E}\eta < \infty$ and $\varepsilon > 0$ is a given number, then $\mathbf{P}\{\eta > \varepsilon\} \leqslant \mathbf{E}\eta/\varepsilon$. Apply this inequality to the components $y_1(t), y_2(t), \ldots, y_n(t)$ of the process $\mathbf{Z}(t)$ using the upper estimates $\widetilde{n}_i(t)$, $i = 1, 2, \ldots, n$ instead of $n_i(t) = \mathbf{E}y_i(t)$.

Introduce the following notations (the symbol T means the transposition operation):

$$\widetilde{m}(t) = (\widetilde{m}_1(t), \widetilde{m}_2(t), \ldots, \widetilde{m}_n(t))^T, \quad \widetilde{n}(t) = (\widetilde{n}_1(t), \widetilde{n}_2(t), \ldots, \widetilde{n}_n(t))^T$$

$$f^*(t) = (f_1^*(t), f_2^*(t), \ldots, f_n^*(t))^T$$

$$\widetilde{m}^{(0)} = (m_1^{(0)}, m_2^{(0)}, \ldots, m_n^{(0)})^T, \quad \widetilde{n}^{(0)} = (n_1^{(0)}, n_2^{(0)}, \ldots, n_n^{(0)})^T$$

$$L = (l_{i,j}), \quad C = (c_{i,j}), \quad D(\widetilde{m}(t-1)) = (d_{i,j}(\widetilde{m}(t-1))$$

$$l_{i,j} = p_{j,i}\rho_j, \quad c_{i,j} = q_{j,i}\gamma_j, \quad d_{i,j}(\widetilde{m}(t-1)) = \gamma_j \ln\frac{1}{1-r_{i,j}}\, p_{i,i}\rho_i \widetilde{m}_i(t-1)$$

$$i, j = 1, 2, \ldots, n, \quad t = 1, 2, \ldots$$

and write system (2.5) in the vector form

$$\begin{aligned} \widetilde{m}(t) &= L\widetilde{m}(t-1) + f^*(t) \\ \widetilde{n}(t) &= (C + D(\widetilde{m}(t-1)))\widetilde{n}(t-1) \\ \widetilde{m}(0) &= \widetilde{m}^{(0)}, \quad \widetilde{n}(0) = \widetilde{n}^{(0)}, \quad t = 1, 2, \ldots . \end{aligned} \tag{3.1}$$

In all subsequent calculations we assume that the inequalities between the vectors from $\mathbf{R}^n$ are considered componentwise, i.e., if $a \in \mathbf{R}^n$, then $a > 0 \Leftrightarrow a_i > 0$, $a \geqslant 0 \Leftrightarrow a_i \geqslant 0$, $i = 1, 2, \ldots, n$.

System (3.1) contains the matrices L, C, D with nonnegative elements. This allows us to apply well-known criteria to the study of the behaviour of $\widetilde{m}(t), \widetilde{n}(t)$ for $t \to \infty$. Describe briefly these criteria. Let the real matrix $H = (h_{i,j})$ have elements $h_{i,j} \geqslant 0$, $i, j = 1, 2, \ldots, n$. All eigenvalues λ_H of this matrix lie in the unit circle if and only if all principal minors of the matrix $I - H$ are positive [2], where I is the identity matrix. Recall also that the matrices H and H^T have the same eigenvalues. Consider a real matrix $S = (s_{ij})$, whose elements are such that $s_{ij} \leqslant 0$ for all $i \neq j$, $i, j = 1, 2, \ldots, n$. The matrix S is called a nondegenerate M-matrix if one of 50 equivalent conditions holds [3]; here we use the following ones: (1) the matrix S^{-1} exists and has nonnegative elements; (2) all principal minors of S are positive; (3) there exists $\xi \in \mathbf{R}^n$, $\xi > 0$, such that $S\xi > 0$. In particular, this implies that $|\lambda_H| < 1$ if and only if $I - H$ is a nondegenerate M-matrix.

Consider the matrices L and C. The structure of the elements of these matrices is such that the following inequalities are valid: $(I - L^T)\xi > 0$, $(I - C^T)\xi > 0$, where $\xi = (1, 1, \ldots, 1)^T$. Therefore, all $|\lambda_L| < 1$, $|\lambda_C| < 1$. Thus, if we consider the matrix

$D(\widetilde{m}(t-1))$ as negligible in system (3.1), then it occurs that $\widetilde{n}(t) \to 0$ for $t \to \infty$. The conditions of the 'smallness' of the matrix $D(\widetilde{m}(t-1))$ are presented in the following assertion.

Proposition 3.1. Let assumptions (H1), (H2) hold for model (1.7) and we additionally assume that there exists $\widetilde{m}^* \in \mathbf{R}^n_+$ such that $I-C-D(\widetilde{m}^*)$ is a nondegenerate M-matrix and there exists $\theta \in (0;+\infty)$ such that $0 \leqslant \widetilde{m}(t) \leqslant \widetilde{m}^*$ for all $t = \theta, \theta+1, \dots$. Then there exists $\lim_{t\to+\infty} \widetilde{n}(t) = 0$ and for each $i = 1,2,\dots,n$ we have $\mathbf{P}\{y_i(t) = 0\} \to 1$ for $t \to \infty$.

Proof. Consider system (3.1) for $t \geqslant \theta$. We can write

$$\begin{gathered} \widetilde{n}(t) = (C + D(\widetilde{m}(t-1)))\,\widetilde{n}(t-1), \quad t = \theta+1, \theta+2, \dots \\ \widetilde{n}(\theta) = \widetilde{n}^{(\theta)} \geqslant 0. \end{gathered} \tag{3.2}$$

Using (3.2), we get

$$0 \leqslant \widetilde{n}(t) \leqslant (C + D(\widetilde{m}^*))^{t-\theta}\, \widetilde{n}^{(\theta)}, \quad t = \theta, \theta+1, \theta+2, \dots$$

which implies that there exists $\lim_{t\to+\infty} \widetilde{n}(t) = 0$. Fix $i = 1,2,\dots,n$ and a number $\varepsilon > 0$. Applying Chebyshev's inequality and the estimate for $n_i(t) = \mathbf{E}y_i(t)$, we get

$$\begin{gathered} 0 \leqslant \mathbf{P}\{y_i(t) > \varepsilon\} \leqslant \frac{\mathbf{E}y_i(t)}{\varepsilon} = \frac{n_i(t)}{\varepsilon} \leqslant \frac{\widetilde{n}_i(t)}{\varepsilon}, \quad t = 1,2,\dots \\ \mathbf{P}\{y_i(t) > \varepsilon\} \to 0, \quad t \to \infty. \end{gathered}$$

These relations complete the proof.

Consider further two cases giving us a chance to get an explicit expression for $\widetilde{m}^*$ in terms of the parameters of the studied problem.

Case 1. Assume $f^*(t) = f^* = \text{const} \geqslant 0$, i.e., $f_i^*(t) = f_i^* = \text{const} \geqslant 0$, $i = 1,2,\dots,n$. System of equations (3.1) with the initial conditions $\widetilde{m}(0) = \widetilde{m}^{(0)}$ has the solution $\widetilde{m}(t) = L^t\,(\widetilde{m}^{(0)} - m_s^*) + m_s^*$, $t = 1,2,\dots$, where $m_s^* = (I-L)^{-1}\,f^* \geqslant 0$ is the stationary solution to this system. Since all $|\lambda_L| < 1$, then $\widetilde{m}(t) \to m_s^*$ for $t \to \infty$ for any initial data. We require that $I-C-D(m_s^*)$ is a nondegenerate M-matrix. Then all principal minors of $I-C-D(m_s^*)$ are positive. Using the continuity of principal minors of the matrix $D(\widetilde{m}(t-1))$ with respect to $\widetilde{m}(t-1)$, one can take $\delta = (\delta_1, \delta_2, \dots, \delta_n)^T$, $\delta_i > 0$, $i = 1,2,\dots,n$ and $\theta \in (0;+\infty)$, so that $m_s^* - \delta \leqslant \widetilde{m}(t) \leqslant m_s^* + \delta$ for all $t = \theta, \theta+1, \dots$, and the principal minors of the matrix $I-C-D(m_s^*+\delta)$ are positive. Therefore, we can assume that $\widetilde{m}^* = m_s^*$.

Case 2. Suppose there exist $\theta \in (0;+\infty)$, $f^* = (f_1^*, f_2^*, \dots, f_n^*) = \text{const}$, $f_i^* = \text{const}$, $i = 1,2,\dots,n$, such that the inequalities $f^*(t) \leqslant f^*$ are valid for all $t = \theta+1, \theta+2, \dots$ and $\widetilde{m}(\theta) \leqslant m_s^* = (I-L)^{-1}\,f^*$, where

$$\widetilde{m}(\theta) = L^\theta\,\widetilde{m}^{(0)} + \sum_{i=0}^{\theta-1} L^i\,f^*(\theta-i).$$

For $t = \theta+1, \theta+2, \dots$, the solution to system of equations (3.1) with the initial conditions $\widetilde{m}(0) = \widetilde{m}^{(0)}$ satisfies the relations

$$\begin{aligned}\widetilde{m}(t) &= L^{t-\theta}\,\widetilde{m}(\theta) + \sum_{i=0}^{t-\theta-1} L^i\, f^*(t-i) \leqslant L^{t-\theta}\,\widetilde{m}(\theta) + \sum_{i=0}^{t-\theta-1} L^i\, f^* \\ &= L^{t-\theta}\,\widetilde{m}(\theta) + \sum_{i=0}^{t-\theta-1} L^i\, f^* + \sum_{i=t-\theta}^{\infty} L^i\, f^* - \sum_{i=t-\theta}^{\infty} L^i\, f^* \qquad (3.3) \\ &= L^{t-\theta}\,\widetilde{m}(\theta) + \sum_{i=0}^{\infty} L^i\, f^* - \sum_{i=t-\theta}^{\infty} L^i\, f^* = L^{t-\theta}\,(\widetilde{m}(\theta) - m_s^*) + m_s^* \leqslant m_s^*.\end{aligned}$$

We have used the formula $\sum_{i=0}^{\infty} L^i = (I-L)^{-1}$ in relations (3.3), which was stipulated by the fact that all $|\lambda_L| < 1$ (the spectral radius of the matrix L is less than one). As the result, we assume $\widetilde{m}^* = m_s^*$.

Now consider certain practical aspects of Proposition 3.1. First of all we note that the inequality $\ln(1/(1-r)) \leqslant r$ holds for all $r \in [0; +\infty)$. For $0 < r \ll 1$ one usually assumes $\ln(1/(1-r)) \approx r$. Therefore, constructing a system of upper estimates for the mathematical expectations $m(t), n(t)$, one can use the functions

$$g_i(z_1, z_2, \dots, z_n) = \sum_{j=1}^{n} \gamma_j\, r_{i,j}\, z_j,\ z_j \geqslant 0, \quad i, j = 1, 2, \dots, n$$

instead of (2.3) and the elements of the matrix $D(\widetilde{m}^*)$ would be the following:

$$d_{i,j}(\widetilde{m}^*) = \gamma_j\, r_{i,j}\, p_{i,i}\, \rho_i\, \widetilde{m}_i^*, \quad i, j = 1, 2, \dots, n, \quad t = 1, 2, \dots \qquad (3.4)$$

Based on (3.4), we calculate the sum of elements of each jth column of the matrix $C + D(\widetilde{m}^*)$ and introduce the set of coefficients

$$R_j(\widetilde{m}^*) = \gamma_j\,(1 + \sum_{i=1}^{n} r_{i,j}\, p_{i,i}\, \rho_i\, \widetilde{m}_i^*), \quad j = 1, 2, \dots, n.$$

For fixed $j = 1, 2, \dots, n$ the coefficient $R_j(\widetilde{m}^*)$ can be interpreted as the mathematical expectation of the number of 'descendants' of an individual from the group B_j in the time period $(t-1; t]$ under the condition that the mean size of the group A_i equals $\widetilde{m}_i^*$, $i = 1, 2, \dots, n$. The coefficient $R_j(\widetilde{m}^*)$ takes into account the probability of survival from the time moment $t-1$ till the moment t for an individual from the group B_j and also the mathematical expectation of the number of individuals from $A_1, A_2, \dots, A_n$ infected by him in the period $(t-1; t]$. Suppose

$$R_1(\widetilde{m}^*) < 1, \quad R_2(\widetilde{m}^*) < 1, \dots, R_n(\widetilde{m}^*) < 1. \qquad (3.5)$$

Considering the matrix $(C + D(\widetilde{m}^*))^T$, from (3.5) we get that the vector $\xi = (1, 1, \dots, 1)^T$ satisfies the inequality $(C + D(\widetilde{m}^*))^T \xi < \xi$, i.e., $I - C - D(\widetilde{m}^*)$ is a nondegenerate M-matrix.

Now consider sufficient conditions for $I-C-D(\widetilde{m}^*)$ to be a nondegenerate M-matrix expressed in terms of the sums of weighted elements of the rows of the matrix $C+D(\widetilde{m}^*)$. Since $I-C$ is a nondegenerate M-matrix, there exists $\xi \in \mathbf{R}^n$, $\xi > 0$, $(I-C)\,\xi > 0$. We require that the elements of the matrix $D(\widetilde{m}^*)$ are such that $(C+D(\widetilde{m}^*))\,\xi < \xi$. For each ith row, $i = 1,2,\ldots,n$, of the matrix $C+D(\widetilde{m}^*)$ we can write the inequality

$$\sum_{j=1}^{n} (c_{i,j} + d_{i,j}(\widetilde{m}^*))\,\xi_j < \xi_i \iff \sum_{j=1}^{n} (c_{i,j} + d_{i,j}(\widetilde{m}^*))\,\xi_{j,i} < 1$$

$$\xi_{j,i} = \frac{\xi_j}{\xi_i}, \quad j = 1,2,\ldots,n.$$

Using the notations for the elements $c_{i,j}$ and relations (3.4), we introduce the coefficients

$$E_i(\widetilde{m}^*) = Q_i(\widetilde{m}^*) + P_i(\widetilde{m}^*), \quad Q_i(\widetilde{m}^*) = \sum_{j=1}^{n} \gamma_j\, q_{j,i}\, \xi_{j,i}$$

$$P_i(\widetilde{m}^*) = p_{i,i}\,\rho_i\,\widetilde{m}_i^* \sum_{j=1}^{n} \gamma_j\, r_{i,j}\, \xi_{j,i}, \quad i = 1,2,\ldots,n.$$

For fixed $i = 1,2,\ldots,n$ the coefficient $E_i(\widetilde{m}^*)$ can be interpreted as the mathematical expectation of the number of 'descendants' of the individuals from the groups $B_1, B_2, \ldots, B_n$, A_i supplementing the group B_i in the time period $(t-1;t]$ with the mean sizes of those groups equal to $\xi_{1,i}, \xi_{2,i}, \ldots, \xi_{n,i}$, $\widetilde{m}_i^*$, respectively. The coefficient $Q_i(\widetilde{m}^*) < 1$ corresponds to the inflow of individuals from the groups $B_1, B_2, \ldots, B_n$ into the group B_i under the mean values $\xi_{1,i}, \xi_{2,i}, \ldots, \xi_{n,i}$ of their sizes; the coefficient $P_i(\widetilde{m}^*)$ denotes the mathematical expectation of the number of individuals of the group A_i infected in the period $(t-1;t]$ under its mean size $\widetilde{m}_i^*$ and the weighed mean $\sum_{j=1}^{n} \gamma_j\, r_{i,j}\, \xi_{j,i}$ of the total number of individuals of the groups $B_1, B_2, \ldots, B_n$. Assuming

$$E_1(\widetilde{m}^*) < 1, \quad E_2(\widetilde{m}^*) < 1, \ldots, E_n(\widetilde{m}^*) < 1 \tag{3.6}$$

we get that $I-C-D(\widetilde{m}^*)$ is a nondegenerate M-matrix.

Thus, within the hypothesis of Proposition 3.1, inequalities (3.5) or (3.6) are sufficient for the extinction of HIV infection within the population. These inequalities are not necessary, but it may occur that one or several inequalities of the form $R_k(\widetilde{m}^*) \geqslant 1$, or $E_l(\widetilde{m}^*) \geqslant 1$ may lead to a violation of the HIV infection extinction condition in a population. Therefore, the groups of individuals B_k, or A_l such that $R_k(\widetilde{m}^*) \geqslant 1$, $E_l(\widetilde{m}^*) \geqslant 1$ may prevent the extinction of HIV infection within a population (under the condition that the sizes of these groups are not identically equal to zero).

4. Numerical experiments with the model

Following [12, 15], consider the groups of HIV-susceptible individuals: A_1, socially adapted individuals, A_2, individuals with a high pathology development risk, A_3, chronic alcoholics, A_4, individuals with drug dependence; and consider also the following HIV-infected individuals: B_1, socially adapted ones, B_2, individuals with a high pathology development risk, B_3, chronic alcoholics, B_4, individuals with drug dependence (see Fig. 1). Denote the mathematical expectations of the total number of susceptible and infected individuals by

$$m_S(t) = \mathbf{E}(x_1(t) + x_2(t) + x_3(t) + x_4(t)), \quad n_I(t) = \mathbf{E}(y_1(t) + y_2(t) + y_3(t) + y_4(t)).$$

The aim of numerical experiments was to study the dynamics of $m_S(t)$ and $n_I(t)$ depending on the parameters of the model, assuming that those parameters satisfy the conditions of Proposition 3.1 (experiment 1) or do not satisfy them (experiment 2). The mathematical expectations $m_S(t)$ and $n_I(t)$ were estimated by standard methods of mathematical statistics based on sample data obtained from 100 implementations of the random process $\mathbf{Z}(t) = (x_1(t), \ldots, x_4(t), y_1(t), \ldots, y_4(t))$, $t = 1, 2, \ldots, T_{\text{mod}}$, where T_{mod} is the duration of the simulated period in days. The duration of the time period $(t-1; t]$ was equal to one day.

In order to obtain implementations of the random process $\mathbf{Z}(t)$ we used the Monte Carlo method. Within a particular implementation, the initial data $x_i^{(0)}$, $y_i^{(0)}$, $i = 1, 2, 3, 4$, are generated for $t = 0$. After that the values $x_i(t-1)$, $y_i(t-1)$, $i = 1, 2, 3, 4$, are fixed for each interval $(t-1; t]$ and then the random variables entering system (1.7) are generated from these values. Then next values of $x_i(t)$, $y_i(t)$, $i = 1, 2, 3, 4$, are calculated for $t = 1, 2, \ldots, T_{\text{mod}}$. Numerical procedures and the random number generator described in [9, 10] were used for the generation of random values entering the equations of the model.

The dynamics of statistical estimates $\bar{m}_S(t)$, $\bar{n}_I(t)$ of the mathematical expectations $m_S(t)$ and $n_I(t)$ is presented in Figs. 2 and 3. Numerical values of the estimates $\bar{m}_S(t)$, $\bar{n}_I(t)$ and the boundaries of the confidence intervals for $m_S(t)$ and $n_I(t)$ are shown in Table 1 for the confidence level $P = 0.95$ and the most typical t.

In the first experiment we used the following values of the parameters of the model. The initial sizes of groups were taken as constants

$$x_1^{(0)} = 294640, \quad x_2^{(0)} = 442614, \quad x_3^{(0)} = 59015, \quad x_4^{(0)} = 25292$$
$$y_1^{(0)} = 0, \quad y_2^{(0)} = 0, \quad y_3^{(0)} = 200000, \quad y_4^{(0)} = 400000.$$

The inflows into the groups A_1 and A_2 in the time period $(t-1; t]$ are given by random variables satisfying the Poisson distribution law with the mathematical expectations $f_1^*(t) \equiv f_1^* = 24.11$, $f_2^*(t) \equiv f_2^* = 34.25$, respectively. The inflows into the groups A_3 and A_4 are absent. The probabilities of survival from $t-1$ till t are the following:

$$\rho_1 = 0.999945, \quad \rho_2 = 0.999932, \quad \rho_3 = 0.999863, \quad \rho_4 = 0.999849$$

$$\gamma_1 = 0.999781, \quad \gamma_2 = 0.999781, quad, \gamma_3 = 0.999753, \quad \gamma_4 = 0.9995891.$$

The probabilities of the transition of individuals in the time period $(t-1;t]$ not related to infection are

$$\begin{gathered}
p_{1,2} = 2.7394 \times 10^{-4}, \quad p_{1,3} = 0, \quad p_{1,4} = 0 \\
p_{2,1} = 1.6437 \times 10^{-4}, \quad p_{2,3} = 0.548 \times 10^{-4}, \quad p_{2,4} = 0.274 \times 10^{-4} \\
p_{3,1} = 0, \quad p_{3,2} = 2.7394 \times 10^{-4}, \quad p_{3,4} = 0 \\
p_{4,1} = 0, \quad p_{4,2} = 3.2871 \times 10^{-4}, \quad p_{4,3} = 0 \\
q_{1,2} = 0, \quad q_{1,3} = 0, \quad q_{1,4} = 0 \\
q_{2,1} = 1.6437 \times 10^{-4}, \quad q_{2,3} = 0, \quad q_{2,4} = 0 \\
q_{3,1} = 0, \quad q_{3,2} = 3.8349 \times 10^{-4}, \quad q_{3,4} = 0 \\
q_{4,1} = 0, \quad q_{4,2} = 4.1087 \times 10^{-4}, \quad q_{4,3} = 0.
\end{gathered}$$

The probabilities of contacts with subsequent infection in the period $(t-1;t]$ are

$$\begin{gathered}
r_{1,1} = 5.4795 \times 10^{-11}, \quad r_{1,2} = 5.48 \times 10^{-11}, \quad r_{1,3} = 0, \; r_{1,4} = 0 \\
r_{2,1} = 5.4795 \times 10^{-11}, \quad r_{2,2} = 5.48 \times 10^{-11} \\
r_{2,3} = 2.74 \times 10^{-11}, \; r_{2,4} = 2.74 \times 10^{-11} \\
r_{3,1} = 0, \; r_{3,2} = 5.48 \times 10^{-11}, \quad r_{3,3} = 8.22 \times 10^{-11}, \quad r_{3,4} = 0 \\
r_{4,1} = 0, \quad r_{4,2} = 8.22 \times 10^{-11}, \quad r_{4,3} = 0, \; r_{4,4} = 1.37 \times 10^{-10}.
\end{gathered}$$

For the second experiment we used

$$\begin{gathered}
y_3^{(0)} = 50000, \quad y_4^{(0)} = 100000 \\
r_{1,1} = 5.4795 \times 10^{-10}, \quad r_{1,2} = 5.48 \times 10^{-10}, \quad r_{1,3} = 0, \; r_{1,4} = 0 \\
r_{2,1} = 5.4795 \times 10^{-10}, \quad r_{2,2} = 5.48 \times 10^{-10} \\
r_{2,3} = 2.74 \times 10^{-10}, \quad r_{2,4} = 2.74 \times 10^{-10} \\
r_{3,1} = 0, \quad r_{3,2} = 5.48 \times 10^{-10}, \quad r_{3,3} = 8.22 \times 10^{-10}, \quad r_{3,4} = 0 \\
r_{4,1} = 0, \quad r_{4,2} = 8.22 \times 10^{-10}, \quad r_{4,3} = 0, \quad r_{4,4} = 1.37 \times 10^{-9}.
\end{gathered}$$

The values of the other model parameters remain the same. Note that the nonzero values of $r_{i,j}$ are increased ten times compared to their values in the first experiment.

The results of the numerical experiments are presented in Table 1 and in Figs. 2 and 3. The parameters of the model in the first experiment are such that

$$\begin{gathered}
R_1(\widetilde{m}^*) = 0.999821, \quad R_2(\widetilde{m}^*) = 0.999827, \quad R_3(\widetilde{m}^*) = 0.999771 \\
R_4(\widetilde{m}^*) = 0.999605.
\end{gathered} \tag{4.1}$$

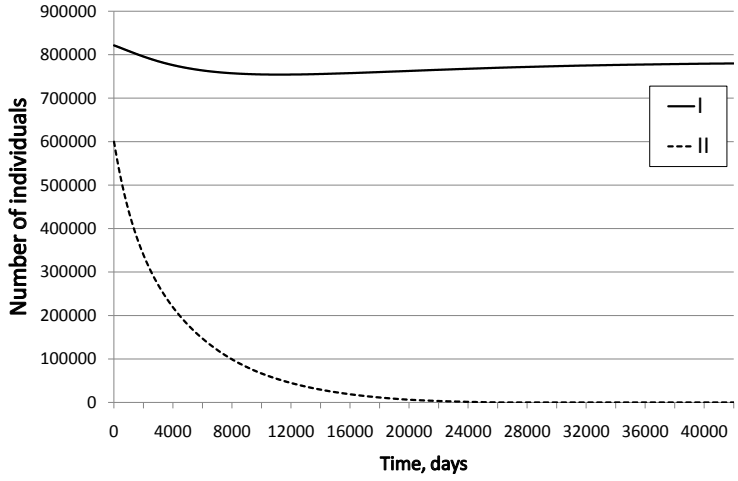

Figure 2. Dynamics of $\bar{m}_S(t)$ (curve I) and $\bar{n}_I(t)$ (curve II) for the first experiment.

Table 1.
Confidence intervals for $m_S(t)$, $n_I(t)$ with the confidence level $P = 0.95$.

Experiment	Value	$t = 0$	$t = 20000$	$t = 40000$
1	$m_S(t)$	821561 ± 0.0	762621 ± 165.914	779341 ± 168.157
	$n_I(t)$	600000 ± 0.0	6272.88 ± 18.111	4.49 ± 0.462
2	$m_S(t)$	821561 ± 0.0	351306 ± 182.967	440861 ± 219.855
	$n_I(t)$	150000 ± 0.0	154431 ± 117.362	91295.7 ± 132.336

Therefore, inequalities (3.5) hold and $I - C - D(\widetilde{m}^*)$ is a nondegenerate M-matrix, which provides the validity of Proposition 3.1. It is seen from Fig. 2 that the infection in the population decays in time in spite of rather large initial sizes of the groups B_3 and B_4.

Proposition 3.1 is not satisfied for the second experiment, because $I - C - D(\widetilde{m}^*)$ is not a nondegenerate M-matrix. Table 1 and Figure 3 show that the mathematical expectation of the number of infected individuals $n_I(t)$ goes to a nonzero stationary level in the course of time. In the second experiment the coefficients $R_i(\widetilde{m}^*)$ have the values

$$R_1(\widetilde{m}^*) = 1.000186, \quad R_2(\widetilde{m}^*) = 1.000239, \quad R_3(\widetilde{m}^*) = 0.999923 \qquad (4.2)$$
$$R_4(\widetilde{m}^*) = 0.999745$$

exceeding (4.1). This excess is caused by the tenfold increase in the contact probabilities compared to the first experiment. It is seen from (4.2) that $R_1(\widetilde{m}^*) > 1$ and $R_2(\widetilde{m}^*) > 1$. Therefore, within the given set of the model parameters, the groups B_1 and B_2 contribute the most to the support of HIV infection. Note that the initial

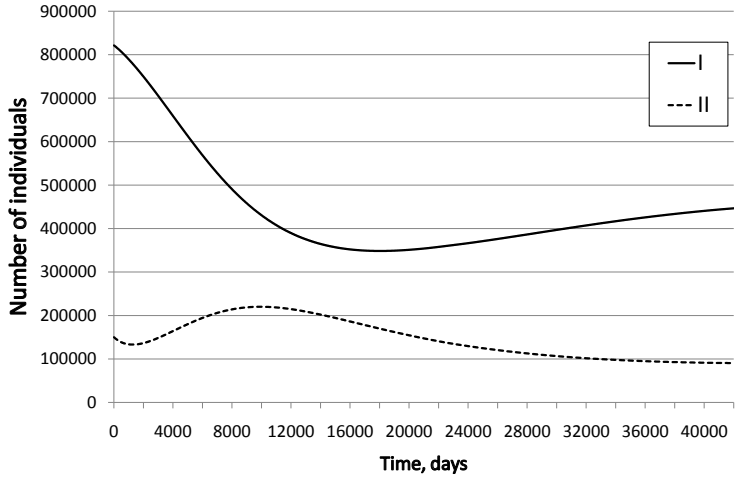

Figure 3. Dynamics of $\bar{m}_S(t)$ (curve I) and $\bar{n}_I(t)$ (curve II) for the second experiment.

sizes of the groups B_1, B_2 are equal to zero, but these groups are supplemented in the course of time because of the inflow of individuals from the groups B_3, B_4, and also of individuals from the groups A_1, A_2. In its turn, the growth of the groups B_1, B_2 increases the flow of infected individuals from the groups A_1, A_2, which increases the total number of infected individuals. The redistribution of the numbers of individuals determines the values of stationary levels for $m_S(t)$, $n_I(t)$.

5. Conclusion

The paper presents a stochastic model of HIV infection spread within a heterogeneous population. The population is considered as a set of groups representing different status of individuals relative to the disease (susceptible, infected, sick, in remission), social stratification, or spatial heterogeneity of the population. Individuals within each group are indistinguishable. The sizes of the groups vary in time as the result of individuals passing from one group to another (because of infection, development of the disease, detection and treatment of sick individuals), because of the group increase by the inflow of new individuals (reaching a particular age or immigration), the outflow of individuals from groups because of natural mortality, emigration to other regions, or death by disease. The dynamics of the sizes of individual groups is described by a system of stochastic differential equations for nonnegative integer-values variables. The dimension of the system of equations is determined by the number of groups used for the description of the population heterogeneity.

The model presented here admits an analytic study and allows us to apply a classification of groups according to the rate of their influence on the decay or

prevalence of the HIV infection in the population. Such classification is related to the hypothesis of Proposition 3.1 and, particularly, to the fulfillment of inequalities (3.5) or (3.6). Considering inequality (3.5), we can easily see that the conditions of Assertion 2 hold if the mean inflows $f_i^*(t)$ of individuals into the groups A_i are sufficiently small, $i = 1,2,\ldots,n$. Therefore, the decrease in the inflows of individuals into the groups $A_1, A_2, \ldots, A_n$ can be considered as a way of eradication of HIV infection in a population.

Note that the inflows $f_i(t)$ of individuals into the groups A_i, $i = 1,2,\ldots,n$, can be nonstationary random processes whose characteristics depend on social-economic conditions. In addition, some variations of the parameters $q_{i,j}$ are possible because of the change of the factors influencing social adaptation of HIV-infected individuals. In this case Proposition 3.1 has an auxiliary character and the study of probabilistic characteristic of the sizes of groups will be conducted by numerical experiments. Based on the results of numerical experiments, one can formulate statistical hypotheses to be later used for analysis and processing of real data related to HIV infection.

References

1. S. Akhtar, T. E. Carpenter, and S. K. Rathi, A chain–binomial model for intra–household spread of Mycobacterium tuberculosis in a low socio–economic setting in Pakistan. *Epidemiol. Infect.* (2007) **135**, No. 1, 27-33.
2. R. Bellman, *Introduction to Matrix Analysis*. McGraw-Hill, 1960.
3. A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, New York, 1979.
4. A. A. Borovkov, *Theory of Probability*. Nauka, Moscow, 1986 (in Russian).
5. D. J. Daley and J. Gani, *Epidemic Modelling: An Introduction*. Cambridge Studies in Mathematical Biology 15. Cambridge Univ. Press, 1999.
6. Global HIV/AIDS response: Epidemic update and health sector progress towards Universal Access. *Progress Report 2011*. WHO/UNAIDS/UNICEF, 2011. http://whqlibdoc.who.int/publications/2011/9789241502986_eng.pdf
7. B. T. Grenfell, O. N. Bjornstad, and B. F. Finkenstadt, Dynamics of measles epidemics: scaling noise, determinism and predictability with the TSIR model. *Ecological Monographs* (2002) **72**, No. 2, 185–202.
8. P. M. Leclerc, A. P. Matthews, and M. L Garenne, Fitting the HIV epidemic in Zambia: A two-sex microsimulation model. *PLos ONE* (2009) **4**, No. 5, e5439.
9. M. A. Marchenko, The complex of programs MONC for distributed computing Monte–Carlo method. *Sib. J. Comp. Math.* (2004) **7**, No. 1, 43–55.
10. G. A. Mikhailov and A. V. Voitishek, *Numerical Statistical Simulation. Monte Carlo Methods*. Akademia, Moscow, 2006 (in Russian).
11. H. Nishiura, Real-time forecasting of an epidemic using a discrete time stochastic model: a case study of pandemic influenza (H1N1–2009). *Biomed. Eng. Online* (2011) **10**, No. 15.
12. E. A. Nosova and A. A. Romanyukha, Regional index of HIV infection risk based on factors of social dysadaptation *Russ. J. Numer. Anal. Math. Modelling* (2009) **24**, No. 4, 325–340.

13. N. V. Pertsev, A. A. Romanyukha, and V. S. Kasatkina, Nonlinear stochastic model of spreading of tuberculosis. *Control Systems Inf. Tech.* (2008), No. 1.2(31), 246–250.

14. N. V. Pertsev and V. N. Leonenko, Stochastic individual–based model of spread of tuberculosis. *Russ. J. Numer. Anal. Math. Modelling* (2009) **24**, No. 4, 341–360.

15. A. A. Romanyukha and E. A. Nosova, Modelling spread of HIV as result of social maladjustment in population. *UBS* (2011) **34**, 227-253.

16. P. M. A. Sloot, S. V. Ivanov, A. V. Boukhanovsky, D. A. M. C. van de Vijver, and C. A. B. Boucher, Stochastic simulation of HIV population dynamics through complex network modelling. *Int. J. Comp. Math.* (2008) **85**, No. 8, 1175–1187.

17. W. Y. Tan and Y. Zhengzheng, Estimation of HIV infection and incubation via state space models. *Math. Biosciences* (2000) **167**, No. 1, 31–50.