

# ИДЕНТИФИКАЦИЯ МАТЕМАТИЧЕСКОЙ МОДЕЛИ РАСПРОСТРАНЕНИЯ КОИНФЕКЦИИ ТУБЕРКУЛЕЗА И ВИЧ В ЭНДЕМИЧНЫХ РАЙОНАХ

В. А. Латышенко<sup>1,2</sup>, В. Н. Каштанова<sup>1</sup>, О. И. Криворотько<sup>1,2</sup>, С. И. Кабанихин<sup>2</sup>

<sup>1</sup>Новосибирский государственный университет, 630090, Новосибирск

<sup>2</sup>Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск

УДК 519.622.2

Теоретически исследована математическая модель распространения коинфекции туберкулеза (ТБ) и ВИЧ. Анализ чувствительности осуществлен с помощью метода собственных значений гессиана матрицы чувствительности. Численные результаты показали, что шесть параметров (из 15 доступных) однозначно идентифицируются по данным измерений количества инфекционных туберкулезом (без ВИЧ), инфицированных ТБ и ВИЧ и больных СПИДом в течение как минимум 3 лет. Анализ чувствительности важно проводить прежде, чем применять численный метод решения возникающей обратной задачи.

**Ключевые слова:** туберкулез, ВИЧ, анализ чувствительности, идентифицируемость, обратная задача.

## Введение

В некоторых регионах России наблюдается «высокая» заболеваемость вирусом иммунодефицита человека (ВИЧ) и туберкулезом (ТБ). Существует высокая вероятность распространения эпидемий этих заболеваний и их коинфекции [1, 2], так как эти болезни влияют друг на друга и имеют сильную связь в период заболевания [3, 4]. Известно, что эпидемия ВИЧ в 1980-х годах вызвала высокий уровень распространения и смертности от ТБ особенно в странах Центральной и Южной Африки. В Российской Федерации за последние десять лет число случаев ВИЧ удвоилось. ВИЧ в России обычно выявляется слишком поздно, в результате чего он перерастает в СПИД. В среднем на 100 000 населения в России насчитывается 335 ВИЧ-инфицированных. Наибольшее количество случаев зарегистрировано в Свердловской области в 2014 году: инфицировано 1000 человек на каждые 100 000 человек [5] и около 70% смертельных исходов от ТБ. Хотя ВИЧ и ТБ могут контролироваться лечением, каждое из этих заболеваний может быть смертельным. Переход от латентного к острому туберкулезу может проходить быстрее при приобретении ВИЧ. С другой стороны, ВИЧ-положительный индивидум более восприимчив к туберкулезу из-за плохой иммунной системы. ТБ считается наиболее распространенной причиной смерти от ВИЧ. Следовательно, обе болезни влияют на распространение друг друга, что делает изучение их коинфекции очень важным.

Математическое моделирование является одним из мощных инструментов для прогнозирования эпидемии коинфекции ВИЧ и ТБ. Но лишь несколько статистических или математических моделей использовались для изучения последствий их совместной динамики на уровне популяции [6]. Большая часть математических моделей описывает одно заболевание и основаны на динамических системах обыкновенных дифференциальных уравнений. Значительное число сосредоточено на ТБ [7–10] или на динамике передачи ВИЧ/СПИДа [11, 12]. В статьях [13–15] имеется несколько моделей коинфекции ТБ/ВИЧ. Naresh и др. [13] разработали нелинейную математическую модель, основанную на четырех нелинейных обыкновенных дифференциальных уравнениях. Их модель была сосредоточена на динамике передачи ВИЧ и ТБ в популяциях разного размера. Porco и др. [14] прогнозируют потенциальное воздействие ВИЧ на вероятность и ожидаемую тяжесть вспышек туберкулеза с использованием моделирования дискретных событий.

---

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации, гранта Президента РФ (номер МК-1214.2017.1) и Министерства образования и науки Республики Казахстан (грант № 1746/ГФ4 «Теория и численные методы решения обратных и некорректных задач естествознания»).

Но большинство из этих моделей были сосредоточены только на структуре моделей, а параметры модели, описывающие характеристики болезни и населения, были усредненными. Поэтому прогнозы распространения эпидемий на основе этих моделей неверны и не учитывают региональные особенности. Следовательно, необходимо качественно оценить параметры модели (или их комбинации) для построения более точной математической модели для конкретной популяции или региона. Эффективным инструментом является подход обратных задач, который заключается в определении или оценке параметров математической модели (коэффициенты динамических систем) с использованием дополнительных измерений о некоторой группе людей (например, ВИЧ-инфицированных, больных СПИДом и т.д.) в фиксированные моменты времени (в годах) [16, 17].

Модель коинфекций ТБ и ВИЧ включает в себя множество известных биологических взаимодействий (параметров) в одну крупную модель. Для больших моделей возникает вопрос о том, могут ли неизвестные параметры однозначно определяться по измеренным данным (анализ чувствительности). Системные подходы к ответу на этот вопрос необходимы, поскольку уникальность значений параметров модели имеет важное значение для интеллектуального моделирования и решения проблемы оценки параметров. В литературе существует несколько типичных методов: метод корреляции, ортогональный метод, метод собственных значений и т.д. Обзор методов приведен в статье [18].

Прежде чем решать проблему оценки параметров, нужно знать, какие параметры можно идентифицировать, а какие нельзя. Идентификация параметров означает, что теоретически возможно однозначно определять значения параметров по доступным данным.

## 1 Математическая модель

Рассмотрим задачу Коши для математической модели распространения коинфекции ТБ и ВИЧ, разработанную группой американских исследователей [6]:

$$\left\{ \begin{array}{l} \frac{dS}{dt} = \Lambda - \beta c S \frac{I+J_3}{N} - \lambda \sigma S \frac{J^*}{R} - \mu S, \\ \frac{dL}{dt} = \beta c (S+T) \frac{I+J_3}{N} - \lambda \sigma L \frac{J^*}{R} - (\mu + k + r_1)L, \\ \frac{dI}{dt} = kL - (\mu + d + r_2)I, \\ \frac{dT}{dt} = r_1L + r_2I - \beta c T \frac{I+J_3}{N} - \lambda \sigma T \frac{J^*}{R} - \mu T, \\ \frac{dJ_1}{dt} = \lambda \sigma (S+T) \frac{J^*}{R} - \beta c J_1 \frac{I+J_3}{N} - (\alpha_1 + \mu)J_1 + r^*J_2, \\ \frac{dJ_2}{dt} = \lambda \sigma L \frac{J^*}{R} + \beta c J_1 \frac{I+J_3}{N} - (\alpha_2 + \mu + k^* + r^*)J_2, \\ \frac{dJ_3}{dt} = k^*J_2 - (\alpha_3 + \mu + d^*)J_3, \\ \frac{dA}{dt} = \alpha_1J_1 + \alpha_2J_2 + \alpha_3J_3 - (\mu + f)A, \\ S(0) = S_0, L(0) = L_0, I(0) = I_0, T(0) = T_0, \\ J_1(0) = J_{10}, J_2(0) = J_{20}, J_3(0) = J_{30}, A(0) = A_0. \end{array} \right. \quad (1)$$

$$N = S + L + I + T + J_1 + J_2 + J_3 + A,$$

$$R = S + L + T + J_1 + J_2, J^* = J_1 + J_2 + J_3.$$

Описание функций и параметров задачи (1) приведены в таблице 1.

Математическая модель основана на следующих предположениях [6]:

- смешение между людьми однородно;
- ВИЧ-инфицированные лица, зараженные туберкулезом ( $J_3$ ), и вирусом СПИД ( $A$ ) не могут получить эффективное лечение ТБ;
- люди заражаются ТБ только через контакты с инфицированными ТБ индивидами ( $I$  и  $J_3$ );
- люди могут заразиться ВИЧ только через контакты с ВИЧ-инфицированными людьми (группа  $J^*$ ).

Таблица 1: Определения параметров и функций, используемых в модели(1)

Параметр	Описание	Единицы	Значение
$\Lambda$	Приток молодежи в модельную популяцию	чел/год	зависит от популяции
$\beta * c$	параметр контагиозности ТБ	—	зависит от популяции
$\lambda * \sigma$	параметр контагиозности ВИЧ	—	зависит от популяции
$\mu$	смертность от причин не связанных с ТБ и ВИЧ	1/год	зависит от популяции
$k$	скорость развития активной формы ТБ (без ВИЧ)	год	0.05
$k^*$	скорость развития активной формы ТБ (с ВИЧ)	год	0.25
$d$	смертность от ТБ	1/год	0.1
$d^*$	смертность от ТБ+ВИЧ	1/год	0.2
$f$	смертность от СПИД	1/год	0.5
$r_1$	скорость лечения от латентного ТБ (без ВИЧ)	чел/год	3
$r_2$	скорость лечения от ТБ (без ВИЧ)	чел/год	1
$r^*$	скорость лечения от латентного ТБ (с ВИЧ)	чел/год	3
$\alpha_1$	скорость перехода ВИЧ к СПИД для группы $J_1$	год	0.1
$\alpha_2$	скорость перехода ВИЧ к СПИД для группы $J_2$	год	0.2
$\alpha_3$	скорость перехода ВИЧ к СПИД для группы $J_3$	год	0.5
$N$	общее количество населения	чел	зависит от популяции
$S(t)$	количество не зараженных индивидов		
$L(t)$	количество индивидов латентно зараженных ТБ (без ВИЧ)		
$I(t)$	количество индивидов с активной формой ТБ (без ВИЧ)		
$T(t)$	количество индивидов вылеченных от ТБ (без ВИЧ)		
$J_1(t)$	количество индивидов зараженных ВИЧ (без ТБ)		
$J_2(t)$	количество индивидов зараженных ВИЧ и латентно зараженных ТБ		
$J_3(t)$	количество индивидов зараженных ВИЧ с активной формой ТБ		
$A(t)$	количество индивидов со СПИД		

Для идентификации параметров модели известна дополнительная информация в фиксированных точках  $t_k$  только о трех группах индивидов:

$$I(t_k) = I_k, \quad J_3(t_k) = J_{3k}, \quad A(t_k) = A_k, \quad k = 1, \dots, K. \quad (2)$$

## 2 Анализ чувствительности

Анализ чувствительности параметров модели распространения коинфекции ТБ и ВИЧ (1) проведен с помощью метода собственных значений [19], который позволяет глобально оценивать и сравнивать влияние значений параметров на данные измерений.

### 2.1 Метод собственных значений

Метод собственных значений находит неидентифицируемые параметры, основываясь на свойствах собственных значений  $\lambda$  и собственных векторов  $u$  матрицы  $H = S^T S$ , где  $S$  это матрица чувствительности, которая определяется следующим образом:

$$S = \begin{pmatrix} s_{11}(t_1) & \dots & s_{1n_\theta}(t_1) \\ \vdots & \ddots & \vdots \\ s_{n_y 1}(t_1) & \dots & s_{n_y n_\theta}(t_1) \\ \vdots & \vdots & \vdots \\ s_{11}(t_K) & \dots & s_{1n_\theta}(t_K) \\ \vdots & \ddots & \vdots \\ s_{n_y 1}(t_K) & \dots & s_{n_y n_\theta}(t_K) \end{pmatrix}. \quad (3)$$

Здесь  $s(t_i)$  определяется по вектору параметров  $\theta^*$ , заданным начальным данным  $S(0), L(0), I(0), T(0), J_1(0), J_2(0), J_3(0), A(0)$  и измерениям  $y_i$  в фиксированные моменты времени  $t_k, k = 1, \dots, K$ , как

$$s_{ij}(t_k) = \frac{\partial y_i(t_k, \theta^*)}{\partial \theta_j} \quad (4)$$

Алгоритм метода собственных значений:

- 1) Задаем критерий остановки  $\delta$ , массив идентифицируемых параметров  $I = 1, \dots, L$  и массив неидентифицируемых параметров  $U = \emptyset$ . Строим матрицу чувствительности  $S$  (3).
- 2) Если  $I$  пусто, то стоп. Для данной модели все параметры неидентифицируемы.
- 3) Считаем матрицу  $H = S^T S$ . Для нее находим собственные значения  $\lambda^l$  ( $\lambda^1 \leq \lambda^l \leq \dots \leq \lambda^L$ ) и собственные вектора  $u^l$ .
- 4) Если  $\lambda^1 \geq \delta$ , то стоп. Параметры, находящиеся в  $I$  идентифицируемы.
- 5) Выбираем  $l: |u_l^1| = \max(|u_1^1|, |u_2^1|, \dots, |u_L^1|)$  (максимальный элемент в собственном векторе, который соответствует минимальному собственному значению). Добавляем  $l$  в множество неидентифицируемых параметров  $U$ , удаляем из  $I$ , удаляем соответствующий столбец из матрицы  $S$  и на шаг 2.

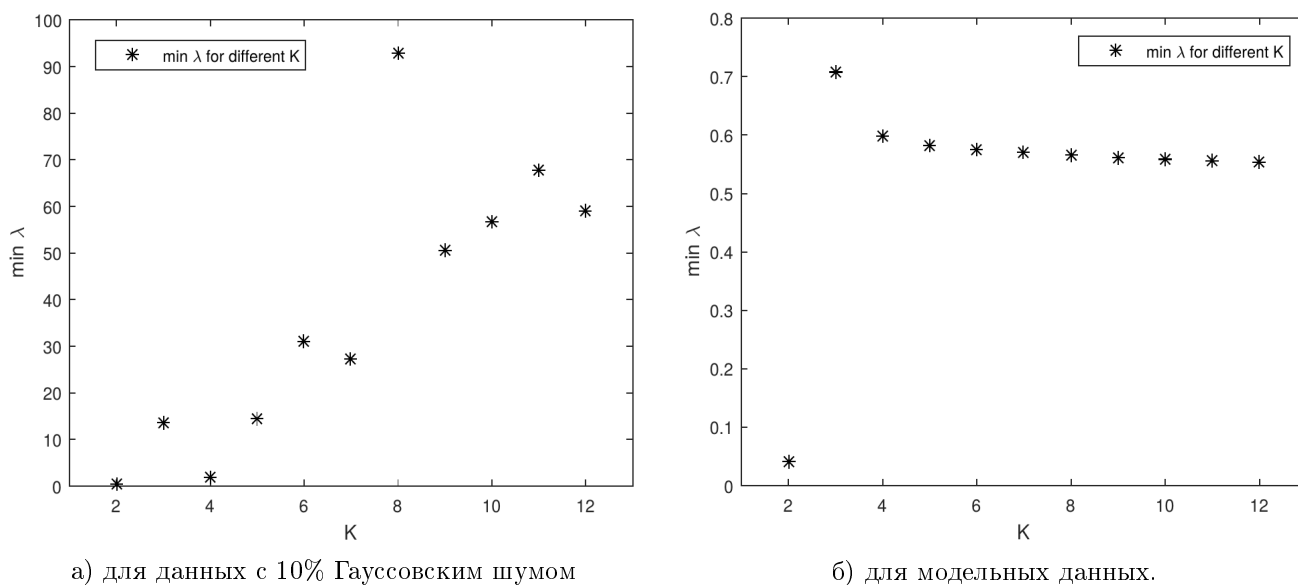
Порядок, в котором параметры удаляются из  $I$ , определяет ранжирование параметров от наименее идентифицируемых до наиболее идентифицируемых.

## 2.2 Результаты

Модель коинфекции ТБ и ВИЧ состоит из уравнений, но мы можем получить информацию  $g(t_k)$  только о трех из них (2) один раз в год в течении, например, 5 лет. Это значит, что мы имеем  $g_i(t_k)$ , где  $i = 1, 2, 3, k = 1, \dots, 5$ . Данная модель содержит множество параметров, некоторые из них можно получить из статической информации. Следовательно, мы проверим модель по 6 параметрам  $\theta^* = (k, k^*, r_2, \alpha_1, \alpha_2, \alpha_3)$ , которые индивидуальны для каждого конкретного региона. В этом случае матрица чувствительности (3) имеем следующий вид:

$$S = \begin{pmatrix} L(t_1) & 0 & -I(t_1) & 0 & 0 & 0 \\ 0 & J_2(t_1) & 0 & 0 & 0 & -J_3(t_1) \\ 0 & 0 & 0 & J_1(t_1) & J_2(t_1) & J_3(t_1) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ L(t_K) & 0 & -I(t_K) & 0 & 0 & 0 \\ 0 & J_2(t_K) & 0 & 0 & 0 & -J_3(t_K) \\ 0 & 0 & 0 & J_1(t_K) & J_2(t_K) & J_3(t_K) \end{pmatrix}$$

Для метода собственных значений 2.1 предполагаем, что  $\varepsilon = 10^{-4}$  [19], значения вектора параметров  $\theta^*$  взяты из таблицы 1 и  $n_\theta = 6$ . После применения алгоритма, получаем  $\min \lambda(H) = 14.03$ , это больше, чем  $\varepsilon$ .



а) для данных с 10% Гауссовским шумом

б) для модельных данных.

Рис. 1: Зависимость между минимальным собственным значением и измерениями

Таблица 2: Минимальное собственное значение для разных  $K$ .

$K$	$\min \lambda$ для данных с ошибкой	$\min \lambda$ для модельных данных
2	0.3094	0.0399
3	13.5082	0.7074
4	1.8698	0.599
5	14.2985	0.5827
6	31.1228	0.576
7	27.3849	0.5707
8	92.9655	0.5659
9	50.4785	0.5616
10	56.5681	0.5583
11	67.6813	0.5558
12	59.047	0.5539

Мы использовали модельные и реальные данные (2) по Свердловской области Российской Федерации с 2010 по 2014 года. Рисунок 1а) показывает зависимость между минимальным собственным значением матрицы  $H$   $\min \lambda(H)$  и количеством измерений  $K$  (одно измерение в год) для данных с 10% Гауссовским шумом и для модельных данных б). Полученные значения приведены в таблице 2. Минимальное собственное значение  $\min \lambda > \varepsilon$  для всех количеств измерений, то есть модель является идентифицируемой по параметрам  $\theta^*$ . Но для  $K = 2$  минимальное собственное значение близко к  $10^{-2}$  для точных данных, эта ситуация может обеспечить практическую идентификацию для другого  $\varepsilon$ . В результате, 6 параметров  $\theta^*$  идентифицируемы для модели (1) с использованием дополнительной статистической информации (2) в точках  $t_k$ ,  $k = 3, \dots, K$  для  $K$  не меньше 3.

### 2.3 Исследование обратной задачи

Так как вектор параметров  $\theta^* = (k, k^*, r_2, \alpha_1, \alpha_2, \alpha_3)$  определяется однозначно, исследуем линеаризованную обратную задачи для модели (1).

Обратная задача (1)–(2) состоит в определении 6 параметров  $\theta^*$  по дополнительным измерениям в фиксированные моменты времени (2). Рассмотрим задачу Коши в векторной форме (1):

$$\dot{x} = y(x(t), \theta)$$

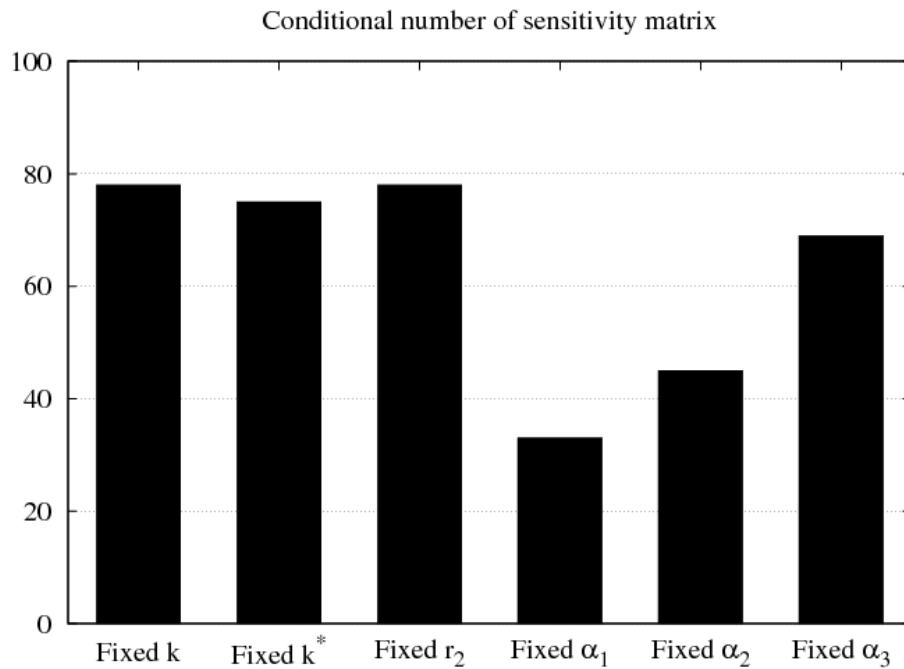


Рис. 2: Число обусловленности матрицы  $S$  для каждого фиксированного параметра из вектора  $\theta^*$ .

и разложим линейную часть в ряд Тейлора в окрестности точки  $\theta^*$  (предполагаем, что  $\theta = \theta^* + q$ , где  $\theta^*$  известный вектор параметров и  $q$  достаточно малое неизвестное отклонение [16]):

$$\dot{x} = y(x(t), \theta^*) + \frac{\partial y(x(t), \theta^*)}{\partial \theta} q + o(|q|^2). \quad (5)$$

Это означает, что в линейном приближении исследование матрицы чувствительности позволяет контролировать некорректность обратной задачи. Действительно, в фиксированных моментах времени уравнение (5) может быть переписано, как:

$$Sq = G. \quad (6)$$

Следовательно,  $G = (\dot{x} - y(x(t), \theta^*))|_{t=t_k}$  описывает известные измерения аналогично (2).

Линеаризованная обратная задача состоит в определении отклонения  $q$  в системе (6). Известно [16], что число обусловленности матрицы  $S$  определяет условность (некорректность) линейной системы (6). Чем больше число обусловленности, тем выше некорректность обратной задачи (для обращения матрицы  $S$  требуется применение методов регуляризации). Число обусловленности матрицы чувствительности удовлетворяет выражению:

$$\text{cond}(S) = \frac{\sigma_{\max}(S)}{\sigma_{\min}(S)}.$$

Здесь  $\sigma_{\max}(S)$  и  $\sigma_{\min}(S)$  максимальное и минимальное сингулярные числа матрицы  $S$ , соответственно.

В данном примере  $\text{cond}(S) = 78$  для задачи идентифицируемости 6 параметров. Это значит, что решение  $q$  рассматриваемой обратной задачи устойчиво к небольшим изменениям данных (уровень шума, вычислительные ошибки и т.д.). Если фиксировать каждый из 6 параметров отдельно  $\theta^*$ , получаем матрицу чувствительности на ранг меньше. Рисунок 2 демонстрирует значения  $\text{cond}(S)$  для каждого фиксированного параметра из  $\theta^* = (k, k^*, r_2, \alpha_1, \alpha_2, \alpha_3)$ . Видно, что число обусловленности в каждом случае невелико, что для исходной ситуации (определение шести параметров) объясняет понятия устойчивости: чем меньше определяемых параметров, тем выше устойчивость. Когда мы фиксируем  $\alpha_1$ , число обусловленности матрицы  $S$  меньше, чем в других случаях. Это значит, что для линейного случая  $\alpha_1$  самый чувствительный параметр в обратной задаче и его нужно контролировать.

## Заключение

Теоретически исследована математическая модель распространения коинфекции ТБ и ВИЧ. Анализ чувствительности осуществляется на основе метода собственных значений для гессиана матрицы чувствительности. Показано, что шесть параметров (из 15 доступных) однозначно идентифицируются по данным измерений количества инфекционных туберкулезом (без ВИЧ), инфицированных ТБ и ВИЧ и больных СПИ-Дом в течение как минимум 3 лет. На практике, если матрица чувствительности имеет полный ранг, но собственные значения разного порядка, параметры, соответствующие наименьшим собственным значениям, теоретически идентифицируемы, но могут быть практически неидентифицируемы. При таких обстоятельствах подходы, основанные на чувствительности, по-прежнему полезны в смысле определения практически неидентифицируемых параметров и управления некорректностью линейаризованной обратной задачи.

## Список литературы

- [1] Эпидемиологические данные о ВИЧ-инфицированных в Российской Федерации. <http://www.hivrussia.ru>.
- [2] Туберкулез в Российской Федерации 2011 г. Аналитический обзор статистических показателей, используемых в Российской Федерации и в мире. — М., 2013. — 280 с.
- [3] S.D. Lawn, A.D. Kerkhoff, M. Vogt and R. Wood. HIV-associated tuberculosis: Relationship between disease severity and the sensitivity of new sputum-based and urine-based diagnostic // BMC Medicine. 2013. V. 11. pp. 231–241.
- [4] P.Nunn, A.Reid and K.M.D. Cock. Tuberculosis and HIV infection: The global setting // Journal of Infections Diseases. 2007. V. 196. pp. 5–14.
- [5] W.H.O. Global tuberculosis report. 2013
- [6] L.W. Roeger, Z. Feng and C. Castillo-Chavez. Modeling tb and HIV co-infections // Mathematical biosciences and engineering. 2009. V. 6, no. 4. pp. 815–837.
- [7] J.P. Aparicio, A.F. Capurro and C. Castillo-Chavez. Markers of disease evolution: the case of tuberculosis // J.Theor. Biol. 2002. V. 215. pp. 227–237.
- [8] S. Blower, P. Small and P. Hopewell. Control strategies for tuberculosis epidemics: New models for old problems // Scitnce. 1996. V. 273. pp. 497–500.
- [9] Z. Feng, C. Castillo-Chavez. Mathematical models for the disease dynamics of tuberculosis, advances in mathematical population dynamics-molecules, cells and man // Vanderbilt University Press. 1998. pp. 117–128.
- [10] T. Porco and S. Blower. Quantifying the intrinsic transmission dynamics of tuberculosis // Theor. Pop. Biol. 1998. V. 54. pp. 117–132.
- [11] H.W. Hethcote. Modeling HIV transmission and AIDS in the United States. Lecture notes in biomathematics ed. Springer-Verlag, New York, 1991.
- [12] H.R. Thieme and C. Castillo-Chavez. How may infection-age infectivity affect the dynamics of HIV/AIDS // SIAM J.Appl.Math. 1993. V. 53. pp. 1447–1479.
- [13] R. Naresh and A. Tripathi. Modeling and analysis HIV-TB co-infection in a variable sized population // Mathematical Modelling and Analysis. 2005. V. 10. pp. 275–286.
- [14] T. Porco, P. Small and S. Blower. Amplification dynamics: Predicting the effect of hiv on tuberculosis outbreaks // Journal of Acquired Immune Deficiency Syndromes. 2001. V. 28. pp. 437–444.
- [15] R.B. Schinazi. Can hiv invade a population which is already sick? // Bull. Braz. Math. Soc.(N.S.). 2003. V. 34. pp. 479–488.

- [16] С.И. Кабанихин. Обратные и некорректные задачи. Новосибирск: Сибирское научное издательство, 2009.
- [17] А.И. Ильин, С.И. Кабанихин, О.И. Криворотько. Об определении параметров моделей, описываемых системами нелинейных дифференциальных уравнений // СЭМИ. 2014. Т. 11. С. 1–14.
- [18] H. Miao, X. Xia, A. S. Perelson and H. Wu. On identifiability of nonlinear ode models and applications in viral dynamics // SIAM Rev Soc Ind Appl Math. 2011. V. 53(1). pp. 3–39.
- [19] T. Quaiser and M. Monnigmann. Systematic identifiability testing for unambiguous mechanistic modeling application to jak-stat, map kinase, and nf-kb signaling pathway models // BMC Systems Biology. 2009. V. 3. no. 50.

*Варвара Александровна Латышенко — магистрант Новосибирского государственного университета, инженер Института вычислительной математики и математической геофизики СО РАН;*  
*e-mail: latushenko\_varia@mail.ru;*

*Виктория Николаевна Каштанова — магистрант Новосибирского государственного университета,*  
*e-mail: vikakashtanova@ya.ru;*

*Ольга Игоревна Криворотько — к.ф.-м.н., н.с. Института вычислительной математики и математической геофизики СО РАН;*  
*e-mail: krivorotko.olya@mail.ru;*

*Сергей Игоревич Кабанихин — член-корр. РАН, директор Института вычислительной математики и математической геофизики СО РАН;*  
*e-mail: kabanikhin@sscc.ru.*