

Применение универсальной математической модели эпидемического процесса «SRID» для прогноза развития эпидемии COVID-19 в городе Москва

Ю.С. Букин^{1,3*}, Ю.П. Джиоев², А.Н. Бондарюк^{3,4}, С.Е. Ткачев⁵, В.И. Злобин²

¹ Лимнологический институт СО РАН, г. Иркутск

² Иркутский государственный медицинский университет МЗ РФ, г. Иркутск

³ Иркутский государственный университет, г. Иркутск

⁴ Иркутский научно-исследовательский противочумный институт Сибири и Дальнего востока, г. Иркутск

⁵ Институт химической биологии и фундаментальной медицины СО РАН, г. Новосибирск

*email: bukinyura@mail.ru

Резюме. В работе представлена универсальная математическая модель эпидемического процесса «SRID», позволяющая с помощью бутстрэп анализа оценить доверительные интервалы для важнейших параметров эпидемического процесса с учетом погрешностей в исходных статистических данных. На основе модели получены варианты прогнозов возможного развития эпидемии COVID-19 в г. Москва по исходным данным за период с 25.03.2020 г. по 21.04.2020 г. при «естественном» течении эпидемического процесса и при принятии эффективных противоэпидемических мер. Значение параметра R_0 (базовое репродуктивное число) в варианте без учета проведения профилактических мер составило от 3,22 до 6,89, что существенно выше этого показателя для других сезонных эпидемий – гриппа и ОРВИ. При таком R_0 пик эпидемии придется на период с 10.05.2020 г. до 05.06.2020 г., а смертность от эпидемии может составить от 24620 до 57032 человек. Для сокращения продолжительности и уменьшения последствий эпидемии необходим комплекс мер, направленных на снижение значения R_0 , к которым следует отнести эффективные изоляционные, карантинные и дезинфекционные мероприятия. В случае их успешной реализации пик эпидемии придется на срок от 10.05.2020 г. до 05.06.2020 г., а смертность по итогам эпидемии может составить от 407 до 2573 человек. Оценка факторов возможного развития эпидемии имеет большое значение для принятия управляющих решений медицинского и экономического характера в масштабах всей РФ. Скрипт и для языка программирования R, реализующие модель «SRID» и первичные данные по г. Москва доступны по ссылке: https://github.com/barnsys/r_srid_model.

Ключевые слова: SARS-CoV-2, COVID-19, математическое моделирования, эпидемиология, SRID.

Application of the universal mathematical model of the epidemic process «SRID» for forecasting the development of the COVID-19 epidemic in Moscow

Yu.S. Bukin^{1,3*}, Y.P. Dzhioev², S.E. A.N. Bondaryuk^{3,4}, S.E. Tkachev⁵, V.I. Zlobin²

¹ Limnological Institute SB RAS, Irkutsk, Russia

² Irkutsk State Medical University, Irkutsk, Russia

³ Irkutsk State University, Irkutsk, Russia

⁴ Irkutsk Antiplague Research Institute of Siberia and Far East, Irkutsk, Russia

⁵ Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia

*email: bukinyura@mail.ru

Summary. The work presents a universal mathematical model of the epidemic process «SRID», which allows using bootstrap analysis to estimate confidence intervals for the most important parameters of the epidemic process, taking into account errors in the initial statistical data. Based on the model, forecast options for the possible development of the COVID-19 epidemic in Moscow were obtained according to the initial data for the period from 25.03.2020 to 21.04.2020 with the «natural» course of the epidemic process and with the adoption of effective anti-epidemic measures. The value of the parameter R_0 (basic reproductive number) in the variant without taking preventive measures ranged from 3.22 to 6.89, which is significantly higher than this indicator for other seasonal epidemics - influenza and SARS. With this R_0 , the peak of the epidemic will be from 10.05.2020 to 05.06.2020, and mortality from the epidemic can be from 24620 to 57032 people. To reduce the duration and reduce the consequences of the epidemic, a set of measures is needed to reduce the R_0 value, which should include effective isolation, quarantine and disinfection measures. If they are successfully implemented, the peak of the epidemic will fall on the period from 10.05.2020 to 05.06.2020, and mortality following the results of the epidemic can be from 407 to 2573 people. Evaluation of the factors of the possible development of the epidemic is of great importance for making management decisions of a medical and economic nature throughout the Russian Federation. The script for the programming language R that implements the «SRID» model and primary data for Moscow is available at: https://github.com/barnsys/r_srid_model.

Key words: SARS-CoV-2, COVID-19, mathematical modeling, epidemiology, SRID.

Введение

В конце 2019 года в г. Ухань, КНР была зафиксирована массовая вспышка острой респираторной инфекции с осложнением в виде вирусной пневмонии с большим количеством летальных исходов. С помощью методов полногеномного анализа было установлено, что эпидемию вызвал новый РНК-содержавший вирус, относящейся к семейству Coronaviridae [1]. Вирус был назван SARS-CoV-2, а вызываемая им тяжелая острая респираторная инфекция получила название COVID-19 [2]. По структуре генома вирус SARS-CoV-2 оказался близок выявленному ранее коронавирусу SARS-CoV – этиологическому агенту эпидемии атипичной пневмонии 2002-2003 годов [1, 3] и вирусу MERS-CoV, вызвавшему вспышку ближневосточного респираторного синдрома 2013 года [1, 4]. Заболевание COVID-19 за короткий период распространилось из Китая практически на все страны мира. 11 марта 2020 года ВОЗ объявила о всемирной пандемии вируса

SARS-CoV-2, требующей принятия безотлагательных мер от правительств всех стран мира. На 10 мая 2020 года в мире насчитывалось около 3,8 миллионов заболевших COVID-19, из этого числа умерло свыше 265 тысяч человек.

В Российской Федерации первые заболевшие COVID-19 стали фиксироваться в конце января 2020 года (приезжие граждане КНР). Массовое распространение инфекции в западных регионах РФ произошло в начале марта 2020 года. Инфекция была занесена гражданами РФ, возвращающимися из стран Европы, охваченных эпидемией COVID-19. В дальнейшем эпидемия COVID-19 в РФ приняла серьезные масштабы (более 209 тысяч зараженных и 1915 умерших к 10 мая 2020 г. Главным очагом распространения вируса SARS-CoV-2 стал город Москва (почти половина всех заболевших и более половины умерших на 21.04.2020). Следует учесть, что г. Москва является главным «хабом» междугородных и международных авиационных, железнодорожных и автомобильных сообщений в России. Вирус SARS-CoV-2 распространяется из этого города по всей территории страны и заносится обратно с территорий других регионов РФ и других стран. Таким образом, ход эпидемического процесса в г. Москва будет определять развитие эпидемии в рамках всей территории РФ. Поэтому чрезвычайно важной задачей, исходя из соображений сохранения здоровья и жизней людей, а также экономических и политических соображений, становится прогнозирование динамики развития, сроков окончания и последствий эпидемии COVID-19 в столице России.

Одним из методов исследования динамики эпидемических процессов является математическое моделирование [5, 6, 7, 8]. Зная параметры эпидемического процесса, (скорость заражения и скорость выздоровления, частоту летальных исходов) можно рассчитать, как после внедрения в популяцию людей или животных кого-либо патогена будут вести себя кривые, демонстрирующие с течением времени количество незараженных, больных и выздоровевших индивидуумов. С помощью этих кривых можно предсказать пики заболеваемости и время окончания эпидемии. Существуют различные классы математических моделей [5, 6, 7, 8], широко применяемые для исследования эпидемий. Класс моделей SRID (Susceptible – восприимчивый, Recovered – выздоровевший, Infectious – инфицированный, Deceased – умерший) предусматривает формирование устойчивого иммунитета к инфекции (повторное заражение невозможно). Класс моделей SIR (Susceptible – восприимчивый, Infectious – инфицированный, Recovered – выздоровевший) – модель с формированием устойчивого иммунитета. Класс моделей SIS (Susceptible – восприимчивый, Recovered – выздоровевший, Susceptible – восприимчивый) – модель без устойчивого иммунитета с возможным хроническим течением болезни. SIS-модели хорошо зарекомендовали себя при описании опасных

вирусных заболеваний с хроническим течением, таких как вирус иммунодефицита человека (HIV) [9, 10], хронические гепатиты В (HBV) [11] и С (HCV) [12]. Модели SIR применяются для описания эпидемических процессов, вызываемых вирусами с относительно легким течением: группы вирусов, вызывающих респираторные инфекции (ОРВИ) [13, 14], и некоторые штаммы вируса гриппа (influenza virus) [15]. Все вышеперечисленные модели могут быть записаны как в виде дифференциальных, так и в виде разностных уравнений. Математические модели позволяют рассчитать такой важный показатель для эпидемического процесса как R_0 – базовое репродуктивное число (индекс репродукции) или количество заболевших, которые будут заражены одним инфицированным за весь период его болезни. Если $R_0 > 1$, эпидемия будет наращивать свой темп, пока не переболеет вся инфицированная популяция, если $R_0 < 1$, то эпидемия быстро пойдет на спад и патоген элиминируется из популяции. Параметр R_0 можно регулировать принятием различных противоэпидемических мер и вакцинацией.

В случае с эпидемией вируса COVID-19 необходимо использование класса моделей SRID. Это связано с достаточно высоким уровнем смертности. В этом отношении COVID-19 похож на недавние вспышки таких вирусных эпидемий как лихорадка Эбола [17], грипп H1N1, ставший причиной пандемии 1918 года («испанка», число жертв – до 50 млн. человек) и 2009 года (свиной грипп, 2627 смертей). В 2002 году произошла эпидемия коронавирусной атипичной пневмонии SARS. Жертвами этого заболевания стали 774 человека. В отличие от лихорадки Эбола, с ярко выраженными клиническими симптомами, заболевание гриппом H1N1 [18] и атипичной пневмонией SARS имеют полиморфную клиническую картину (от бессимптомного течения до тяжелого заболевания со смертельным исходом). Поэтому выявить их эпидемиологические закономерности, основываясь на данных статистики, практически невозможно. Для описания динамики эпидемического процесса этих вирусов активно использовались SRID-модели [19, 20, 21]. Результаты расчетов по этим моделям позволили оценить эпидемические параметры, определяющиеся биологическими свойствами вирусов – возбудителей гриппа H1N1 и SARS. Наиболее важным результатом стали оценки реальных масштабов эпидемий – определение числа людей, переболевших в легкой форме и оказавшихся за пределами статистики тех, кому был поставлен точный диагноз. Математические расчеты показывают, что реальное количество переболевших в ходе пандемии гриппа H1N1 и атипичной пневмонии SARS-CoV намного больше, чем реальное количество зарегистрированных случаев заражения. В этом заключается сложность математических моделей и для эпидемии вируса SARS-CoV-2. Поэтому при прогнозировании характера эпидемии будут значительные погрешности, связанные с

недооценкой количества зараженных и выздоровевших [22]. Официальная статистика выдает более надежные данные только по количеству умерших.

Уже выпущена серия работ, посвященных математическому моделированию эпидемии вируса SARS-CoV-2 [23, 24, 26, 26, 27, 28]. Представленные математические модели основываться на стохастическом подходе имитационного моделирования, на описании процесса с помощью систем дифференциальных уравнений и на системах разностных уравнений. Эти математические модели для оценки параметров использовали реальные данные по развитию эпидемического процесса в Китае и Италии – стране с наиболее напряженным течением эпидемии. Предложенные в работах методы, позволяют оценить статистические доверительные интервалы для кривых протекания эпидемии, параметров эпидемического процесса и значения базового репродуктивного числа R_0 . Результаты моделирования, представленные в работе [26], основываясь на данных работы [22], говорят о том, что вероятней всего общее число инфицированных превышает в 20 раз данные официальной статистики, а число переболевших и выздоровевших недооценено в 40 раз.

Целью работы является построение универсальной математической модели, позволяющей на основе данных о начальных этапах эпидемии дать прогноз ее развития с учетом статистических ошибок. Модель с оценкой доверительных интервалов должна предсказывать количество умерших в течение эпидемии, дату пика заболеваемости, дату окончания эпидемии и значение эффективного репродуктивного числа R_0 . Также модель должна давать оценки прогнозов эпидемии в случае принятия срочных противоэпидемических мер. С учетом разработанной модели сделать возможные прогнозы развития эпидемии COVID-19 для города Москва, основываясь на статистических данных за период с 25.03.2020 по 21.04.2020.

Материалы и методы

Описание модели. Для моделирования эпидемического процесса мы использовали классический подход SRID для эпидемиологических моделей с дискретным временем в предположении формирования устойчивого иммунитета у переболевших [5, 6, 7, 8, 26]. Предполагается, что процесс распространения инфекции происходит в популяции людей с общей численностью N человек. Обозначим: S – количество человек из популяции которые еще не проконтактировали с патогеном; R – количество выздоровевших людей, которые проконтактировали с патогеном и получили стойкий иммунитет; I – количество болеющих или вирусоносителей; D – количество умерших; K – общее количество инфицированных и выздоровевших ($K=I+R+D$, $K=N-S$); Z – количество новых заражений

в течении суток (скорость прироста числа инфицированных, скорость прироста K). Переменные в модели должны удовлетворять условиям $N=S+I+R+D$, $N=S+K$. Предположим, что шаг изменения времени t в модели равен одним суткам (1-е, 2-е, 3-ьи сутки и т. д.). Момент времени $t=1$ – это время появления в популяции первого заболевшего (нулевого пациента). Разностная система уравнений такой модели запишется следующим образом:

$$\left\{ \begin{array}{l} S(t) = S(t-1) - \frac{\alpha}{N} S(t-1) I(t-1) \\ I(t) = I(t-1) + \frac{\alpha}{N} S(t-1) I(t-1) - \beta I(t-1) - \gamma I(t-1) \\ R(t) = R(t-1) + \beta I(t-1) \\ D(t) = D(t-1) + \gamma I(t-1) \\ Z(t) = \frac{\alpha}{N} S(t-1) I(t-1) \\ K(t) = N - S(t) \end{array} \right. \quad (1)$$

В этой системе уравнений фигурируют три свободных параметра: α – скорость заражения в единицу времени (количество зараженных одним больным других людей за одни сутки); β – скорость выздоровления в единицу времени; γ – скорость смерти в единицу времени. Значения параметра α определяется заразностью вируса (биологические свойства патогенна, определяющие вероятность заражения при непосредственном контакте больного и здорового, устойчивостью патогенна в окружающее стере и т. п.) и количеством контактов инфицированного с другими здоровыми людьми (скоростью перемешивания людей в популяции). Значения параметров β и γ определяются биологическими свойствами патогена и эффективностью терапии, при возникновении заболевания.

Начальные условия для модели при старте с момента времени $t=1$ будут следующими: $S(1)=N-1$, $I(1)=1$, $R(1)=0$, $D(1)=0$, $Z(1)=\alpha/N$ и $K(1)=1$. При старте с других моментов времени берутся начальные значения переменных, соответствующих этому моменту времени.

Алгоритм расчета параметров на основе реальных данных. Если мы наблюдаем начальный этап развития эпидемии некоторого инфекционного заболевания в популяции из N человек, то у нас имеется ряд данных, описывающих значение переменных S , R , K , D и I в течение ряда дней развития процесса. Пусть td – момент времени, до которого нам известны реальные данные о развитии эпидемического процесса. Тогда время t будет меняться от 1 до td ($t = 1, 2, 3, \dots, td$). Обозначим реальные данные об эпидемическом процессе за время td как S' , R' , I' и D' . Исходя из данных по изменению реальных значений переменных на каждом временном шаге и уравнения (1), можно рассчитать выборки значений параметров α_t , β_t и γ_t по формулам:

$$\alpha_t = -\frac{N(S'(t) - S'(t-1))}{S'(t-1)I'(t-1)} \quad \bar{\alpha} = \frac{1}{td-1} \sum_{t=2}^{td} \alpha_t \quad (2)$$

$$\beta_t = \frac{R'(t) - R'(t-1)}{I'(t-1)} \quad \bar{\beta} = \frac{1}{td-1} \sum_{t=2}^{td} \beta_t \quad (3)$$

$$\gamma_t = \frac{D'(t) - D'(t-1)}{I'(t-1)} \quad \bar{\gamma} = \frac{1}{td-1} \sum_{t=2}^{td} \gamma_t \quad (4)$$

в этих выражениях $\bar{\alpha}$, $\bar{\beta}$ и $\bar{\gamma}$ – средние значения параметров, которые можно использовать в уравнении (1) как оцененные значения параметров α , β и γ для прогноза эпидемического процесса.

При вычислениях по формулам (2), (3) и (4) мы получаем выборки значений α_t , β_t и γ_t , содержащие $td-1$ элементов. Эти выборки можно использовать для вычисления доверительных интервалов оцененных значений параметров α , β и γ .

Часто такая статистка значений переменных S' , R' , I' и D' может вестись не с момента времени появления нулевого пациента или статистка начального периода эпидемии может быть недостоверной. Тогда оценки параметров α , β и γ можно проводить не с первого дня начала эпидемии, а с момента времени tn , когда значения переменных S' , R' , I' и D' более обоснованы. В формулах (2), (3) и (4) t начальное будет равно не 2, а $tn+1$.

Параметр R_0 – эффективное репродуктивное число будет рассчитываться по формуле:

$$R_{0t} = \frac{\alpha_t}{\beta_t + \gamma_t} \quad \bar{R}_0 = \frac{\bar{\alpha}}{\bar{\beta} + \bar{\gamma}} \quad (5)$$

Здесь R_{0t} – выборка значений базового репродуктивного числа, \bar{R}_0 – среднее значение R_0 .

Оценка доверительных интервалов для переменных при прогнозных расчетах. Так как в силу случайных обстоятельств поведения реальных значений параметров S' , R' , I' и D' будут иметь статистический разброс и значения параметров α , β и γ оцениваются с некоторой статической ошибкой. В этом случае мы должны говорить не об истинных кривых $S(t)$, $I(t)$, $R(t)$, $D(t)$, $K(t)$ и $Z(t)$, а о семействах кривых, дающих разные прогнозы развития эпидемического процесса. Задача состоит в определении доверительного коридора, в пределах которого с выбранной доверительной вероятностью (например, с 95% доверительной вероятностью) будут проходить все кривые, описывающие поведение переменных $S(t)$, $I(t)$, $R(t)$, $D(t)$, $K(t)$ и $Z(t)$. Для оценки доверительного коридора можно использовать непараметрический бутстрэп анализ [29]. Мы предлагаем использовать бутстрэп анализ по следующей схеме: 1) для выборок, рассчитанных по формулам (2), (3), (4) и (5) значений α_t , β_t , γ_t и R_{0t} , делаем необходимое

количество бутстреп реплик; 2) для каждой бутстреп реплики рассчитываем значение $\bar{\alpha}$, $\bar{\beta}$, $\bar{\gamma}$ и \bar{R}_0 (если генерировалось n реплик, то мы будем иметь серию из n оцененных значений $\bar{\alpha}$, $\bar{\beta}$, $\bar{\gamma}$ и \bar{R}_0); 3) для каждой комбинации из n значений параметров проводим расчеты по уравнению (1) в течение исследуемого промежутка времени t ; 4) для каждого момента времени находим диапазоны значений $S(t)$, $I(t)$, $R(t)$, $D(t)$, $K(t)$ и $Z(t)$, между которыми с заданной доверительной вероятностью будет проходить одна из кривых, описывающих поведение оцениваемых исследуемых переменных.

Бутстреп метод в такой реализации позволяет оценить доверительный интервал для даты прохождения пика заболеваемости и для общего количества умерших после окончания эпидемии. Пик заболеваемости будет детектироваться значением t , когда кривая $I(t)$ будет иметь максимальное значение. При расчетах со значениями параметров $\bar{\alpha}$, $\bar{\beta}$ и $\bar{\gamma}$ из бутстреп реплик мы будем иметь серию кривых $I(t)$ с максимальными значениями в разные моменты времени t . Таким образом, мы получим выборку бутстреп значений моментов времени t пика заболеваемости. По этой выборке можно определить доверительный интервал даты прохождения пика заболеваемости. В конечный момент времени, когда число болеющих $I(t)$ будет равно 0 (или близко к 0) мы будем иметь бутстреп реплику значений переменной $D(t)$ по которой можно оценить доверительный интервал для предполагаемого количества умерших.

Расчеты по модели с реализацией бутстреп анализа были проведены с помощью алгоритма, реализованного на языке программирования R. Бутстреп реплики для параметров α , β , γ и R_0 генерировались с помощью пакета «boot» для R (1000 реплик). Для всех параметров и переменных оценивались 95% доверительные интервалы. Скрипт для R, реализующий всю последовательность расчетов, доступен для свободного некоммерческого использования по ссылке: https://github.com/barnsys/r_srid_model.

Массив статистических данных по г. Москва. Для прогнозирования хода эпидемического процесса COVID-19 в г. Москва исходные данные о течении эпидемии были взяты с сайта <https://coronavirus-monitor.ru>. Использовали данные от момента 25.03.2020 г., когда данные по эпидемическому процессу стали включать в статистику по общему количеству инфицированных $K'(t)$, по количеству выздоровевших $R'(t)$ и количеству умерших $D'(t)$. Конечной точкой для оценки параметров было взято 21.04.2020 г. (момент проведения расчетов по модели). Таким образом, оценки значений параметров α , β и γ проводились на основе данных о 28 днях развития эпидемического процесса. Таблицу исходных данных для алгоритма на языке R можно получить по той же ссылке: https://github.com/barnsys/r_srid_model.

Результаты

Расчет параметров модели. Как указано в работе [26] возможно два предположения (сценария развития событий) для проведения расчетов. В первом предположении (сценарии) считается, что мы имеем информацию об истинных значениях количества инфицированных $K'(t)$, количества выздоровевших $R'(t)$ и количества умерших $D'(t)$. Во втором предположении (сценарии) истинное количество инфицированных в 20 раз превышает выдаваемые статистикой данные, а количество выздоровевших в 40 раз превышает выдаваемые статистикой данные. $K'(t)$ умножается на 20, $R'(t)$ умножается на 40. Во втором предположении количество умерших $D'(t)$ не меняется. Исходя из первого предположения, были получены следующие значения параметров: α равнялось 0,184 с доверительным интервалом (0,148 – 0,242); $\beta = 0,0106$ с доверительным интервалом (0,0076 – 0,0141); γ равнялось 0,0013 с доверительным интервалом (0,000932 – 0,00176). Исходя из второго предположения, α равнялось 0,196 с доверительным интервалом (0,159 – 0,253); $\beta = 0,0212$ с доверительным интервалом (0,0151 – 0,0282); γ равнялось 0,00006529 с доверительным интервалом (0,00004661 – 0,0000876). Если мы предполагаем второй сценарий развития событий, то это значительно увеличивает скорости заражения и выздоровления α и β , и уменьшение скорости смерти γ почти в 20 раз.

Для первого сценария параметр R_0 – базового репродуктивного числа равнялся 8,69 с доверительным интервалом (5,85 – 13,72). Для второго сценария R_0 равнялось 4,59 с доверительным интервалом (3,22 – 6,89).

Соответствие истинных кривых $K'(t)$, $R'(t)$, $D'(t)$ и расчетных кривых $K(t)$, $R(t)$, $D(t)$, полученных с вычисленными значениями параметров, за период развития эпидемического процесса с 25.03.2020 г. по 21.04.2020 г. приведено на рисунке 1. На рисунке видно, что данные реальных наблюдений для переменных количества инфицированных $K'(t)$ и количества умерших $D'(t)$ практически точно лежат на расчетные кривые $K(t)$ и $D(t)$. Реальные данные по выздоровевшим $R'(t)$ немного отклоняются от расчетной кривой $R(t)$, но при этом всегда укладываются в рамки доверительного интервала для $R(t)$. Можно сделать вывод, что на начальном временном отрезке развития эпидемического процесса в г. Москва, поведение переменных четко укладывается в рамки модели SRID, описываемой уравнением (1). Используемая математическая модель позволит получить достаточно точный прогноз (в рамках статистики доверительных интервалов) развития эпидемического процесса в рамках двух рассматриваемых сценариев.

Прогноз сценариев развития эпидемии. При расчетах по двум сценариям развития эпидемического процесса с параметрами, оцененными по реальным значениям $K'(t)$, $R'(t)$ и скорректированными в сторону увеличения значениям $K'(t)$, $R'(t)$ динамика процесса оценивалась от 1-го (15.03.2020) дня с момента появления в г. Москва первого заболевшего до 200-го дня (30.09.2020). По этим двум сценариям к 200-му дню эпидемии переболеют все жители Москвы и число людей $S(t)$, которые не проконтактировали с патогенном и не получили иммунитета будет равно 0 (верхняя и нижняя граница доверительных интервалов также будет равна 0). Динамика развития эпидемического процесса по двум сценариям приведена на рисунке 2.

Анализ кривых рисунка 2 показывает, что по первому сценарию в случае нескорректированных значений параметров α , β и γ уравнение (1) показывает более продолжительную по времени эпидемию. Переменная $S(t)$ – количество человек из популяции, которые еще не проконтактировали с патогенном, выйдет в нулевые значения в районе 150-го дня эпидемического процесса. Пик заболеваемости (максимальное значение заболевших и инфицированных $I(t)$) придется на период в районе 100-го дня эпидемии, затем число больных начнет сокращаться. К моменту окончания эпидемии количество умерших достигнет в среднем одного миллиона человек. Этот сценарий является самым пессимистическим, но как показывает опыт других стран (КНР, Южная Корея, Япония, Италия, Испания, США) маловероятным. Более вероятным является второй сценарий, при котором переменная $S(t)$ выйдет в нулевые значения приблизительно на 80-й день эпидемии. Пик заболеваемости придется на 75 день эпидемии. По этому сценарию количество умерших будет значительно меньше (см. рис. 2 – 2г) и составит приблизительно 37000 человек.

Наиболее важными характеристиками эпидемического процесса является дата пика заболеваемости и количество умерших после завершения эпидемии. Наша математическая модель позволяет оценить эти цифры с определением доверительных интервалов. Результаты оценок представлены на рисунке 3. Если развития событий идут по первому сценарию (параметры модели вычислены без коррекции наблюдаемых значений числа инфицированных и выздоровевших), то пик заболеваемости в г. Москва с 95% вероятностью придется на период с 75-го по 112-й день эпидемии (с 28.05.2020 г. по 04.07.2020 г.). Число умерших при этом с 95% вероятностью будет колебаться в интервале от 771644 до 1541847 человек. Если события будут развиваться по второму сценарию (параметры модели вычислены с учетом коррекции наблюдаемых значений числа инфицированных и выздоровевших), то пик заболеваемости с 95% вероятностью придется на период с 57-го по 83-й день эпидемии (с 10.05.2020 г. по 05.06.2020 г.). Число умерших

при втором сценарии с 95% вероятностью будет лежать в диапазоне от 24620 до 57032. Ориентироваться необходимо на второй сценарий развития эпидемии.

Прогноз сценариев развития эпидемии с учетом противоэпидемических мероприятий. В рамках модели SRID, описываемой уравнением (1), изменение значений параметров α – скорость заражения в единицу времени и R_0 – базовое репродуктивное число (см. уравнение 5) будет менять ход эпидемического процесса. Величина параметра α зависит от вероятности заражения при непосредственном контакте с патогенным и от количества контактов больного или вирусоносителя со здоровыми людьми. Все эти показатели могут быть отрегулированы комплексом противоэпидемических мероприятий, снизив, таким образом, значение α и R_0 .

С помощью модели, описываемой уравнением (1) мы сделали расчеты в предположении, что предпринятые противоэпидемические меры сработают, и величина параметра α , начиная от момента 21.04.2020 г., резко сократится. Для расчетов мы использовали параметры модели, полученные для второго более вероятного сценария эпидемического процесса, когда число инфицированных в 20 раз, а выздоровевших в 40 раз превышает выдаваемые статистикой данные. В результате было установлено, что если значение α снизится в 55 раз от текущего (текущее среднее значение равно 0,196), то переменные в модели резко меняют свои траектории (см. рис. 4). Базовое репродуктивное число R_0 при этом уменьшится до 0,083 (благоприятное значение, значительно меньшее 1). Общее количество инфицированных $K(t)$ выходит на плато, количество выздоровевших с течением времени постепенно увеличивается, скорость прироста числа заболевших за сутки $Z(t)$ резко сокращается практически до 0. Траектории кривых $K(t)$, $Z(t)$, и $R(t)$, при таком развитии эпидемического процесса, напоминают развитие событий в таких странах как Китай (КНР) и Южная Корея (РК).

При срабатывании предпринятых противоэпидемических мер сценарий развития событий в г. Москва приведет к тому, что количество умерших $D(t)$ сократиться до 925 человек с 95% доверительным интервалом от 407 до 2573. Общее число переболевших $K(t)$ сократится до 589133 человек с 95% доверительным интервалом от 271904 до 1576192. Как мы видим, масштабы потерь будут существенно меньше. Следует заметить, что такое переключение траекторий переменных происходит только при уменьшении параметра α в 55 раз и более от текущего значения. Если α уменьшится в меньших масштабах, то эффект от противоэпидемических мероприятий будет незначительным. Повторится картина развития событий по второму сценарию, но более растянутая во времени.

Обсуждение

Мы в своих расчетах сделали прогноз динамики эпидемии в г. Москва по двум сценариям: без увеличения оценок количества инфицированных $K'(t)$, количества выздоровевших $R'(t)$ (первый сценарий эпидемии) и с учетом этих поправок (второй сценарий эпидемии). Необходимо понять, какой из сценариев наиболее вероятен для нынешней ситуации развития событий. Прогноз по уровню смертности в двух сравниваемых сценариях сильно отличается (первый сценарий в среднем дает почти в 29 раз больше умерших (см. рис. 3). Статистика по другим странам (Китай, Япония, Южная Корея), в которых заболеваемость пошла на снижение, говорит о том, что первый сценарий в г. Москва не реализуется. Кроме этого (23-24 марта 2020 г.) в новостных лентах (<https://www.livescience.com/covid-antibody-test-results-new-york-test.html>, <https://www.france24.com/en/20200423-fifth-of-new-yorkers-infected-with-covid-19-antibody-test-suggests>) было опубликовано заявление мэра города Нью-Йорка, в котором были обнародованы результаты массового тестирования жителей, на антитела к COVID-19. Осуществлено тестирование 3000 тысяч жителей, у которых не было симптомов заболевания. Исследование показало, что положительные тесты составили 14%. Это означает, что около 2,6 миллионов жителей Нью-Йорка уже переболели COVID-19, большая часть которых не попала в списки зараженных по официальной статистике. Результаты этого анализа показывают, что реальные масштабы эпидемии недоучтены. Поэтому все расчеты и прогнозы необходимо делать с поправками на увеличение количества заболевших и выздоровевших. Данные, выдаваемые статистикой, необходимо корректировать, число инфицированных надо умножать на 20, а количество выздоровевших на 40 как рекомендуется в работе [26].

Число умерших, при втором сценарии с корректировкой числа заболевших и выздоровевших будет лежать в диапазоне 24620 до 57032. По данным Росстата (<https://rosinfostat.ru/smertnost/>) в центральном федеральном округе в 2019 году смертность составляет 12,9 человек на 1000 человек в год. При таком уровне смертности в г. Москва за 150 дней должно умереть 67287 человек. Эпидемия COVID-19 может дать прибавку от 36% до 84% к этой цифре, что является существенным фактором. В 2018 году в России от инфекционных заболеваний умерло 34626 человек. В пересчете на население г. Москва число умерших от инфекционных заболеваний в 2018 году оставило приблизительно 2994 человек. 150 дней эпидемия COVID-19 могут увеличить число умерших от инфекционных заболеваний за один год в г. Москва от 8 до 19 раз. Прогноз, представленный в нашей работе, показывает, что, если не предпринять срочных противоэпидемических мер, то при наиболее благоприятном развитии событий жертвами эпидемии COVID-19 станет значительно больше людей, чем умрет от остальных инфекционных заболеваний вместе

взятых. Это следует подчеркнуть особо, в свете постоянно высказываемых в средствах массовой информации мнений о преувеличенной опасности COVID-19.

Рассчитанное, с помощью модели, значение параметра базового репродуктивного числа R_0 , показывает, что один инфицированный в г. Москва заразит в среднем от 3.22 до 6,86 человек. Эти цифры перекликаются с данными по эпидемии в г. Ухань, КНР, где значение R_0 лежало в пределах цифр 1,4-5,7, рассчитанных для начальной фазы эпидемии, до принятия экстренных карантинных мер [30, 31, 32]. Значение R_0 , для COVID-19 в г. Москва, превышает показатели для сезонных эпидемии гриппа, где R_0 меняется от 0,9 до 2,1 [19] и серотипа гриппа АН1N1 [33] (R_0 лежало в пределах 1.4 - 2 для эпидемии 1918 г. и в пределах 1.15 – 1.90 для эпидемии 2009 г.).

В рамках борьбы с эпидемией COVID-19 власти в г. Москва принимают экстренные меры. Одной из таких мер является введение режима самоизоляции, перевод сотрудников организаций (школы, вузы) на дистанционную работу и учебу, запрет массовых мероприятий, закрытие торговых центров, увеселительных заведений. Другой комплекс мер направлен на уменьшение вероятности заражения, даже если непосредственный контакт больного и здорового человека произошел. К этим мерам относятся ношение защитных масок, дезинфекция помещений, фасадов домов, тротуаров, частое мытье и дезинфекция рук. Представленный в работе прогноз развития эпидемии COVID-19 показывает, если удастся значительно сократить параметр R_0 (через сокращения параметра α - скорость заражения в единицу времени (см. уравнение 1 и 5), то эпидемия в течение относительно короткого промежутка времени сойдет на нет (см. рис. 4). Мероприятия должны быть такого масштаба, чтобы сократить значение R_0 в 55 раз. При таком развитии событий смертность в г. Москва по итогам эпидемии составит от 407 до 2573. Так как, согласно математической модели, большая часть болеющих не выявлена и не изолирована от окружающих, особую значимость приобретают самоизоляция и, возможно, принудительная изоляция части населения, а также комплекс мероприятий по дезинфекции. Вероятно, на момент времени 21.04.2020 г. в таких странах как Китай, Япония, Южная Корея, где развитие эпидемии пошло по сценарию с резким падением заболеваемости (см. рис. 4 часть 2), проводимый комплекс противоэпидемических мероприятий привел к резкому сокращению параметра скорости заражения и «фазовому переходу» в динамике эпидемического процесса.

Результат исследования, связанный с оценкой времени наступления пика заболеваемости с 10.05.2020 г. по 05.06.2020 г. в г. Москва (пик заболеваемости по второму, благоприятному сценарию развития эпидемии), является важной информацией для планирования противоэпидемических мероприятий. Указанные оценки могут быть

использованы при составлении прогнозов ущерба, связанных с простоем экономики. Ориентируясь на эти данные можно определить дату прекращения ограничительных мероприятий. Также следует сказать, что при возможном резком снижении из-за принимаемых мер эффективной скорости заражения, весь эпидемический процесс закончится гораздо быстрее.

Следует сказать, что представленная здесь модель не учитывает того, что периодические вспышки ОРВИ, в том числе и коронавирусной природы, носят сезонные характер. Как показывают предварительные исследования, скорость распространения COVID-19 не зависит от температуры воздуха [34], но исключать климатический и сезонный факторы пока полностью нельзя. Еще одним обстоятельством, неучтенным в работе, которое может затормозить эпидемию COVID-19, это возможность получения в ближайшее время эффективной вакцины.

Представленный в работе алгоритм, реализованный на языке программирования R, является универсальным и позволяет делать прогноз, исходя из начальных данных о развитии эпидемического процесса для любых инфекционных заболеваний, динамика эпидемий которых соответствует модели SRID, в случае, когда к патогену может формироваться длительный иммунитет. Преимуществом его является встроенная статистическая оценка доверительных интервалов для всех исследуемых переменных и параметров эпидемического процесса.

В результате проведенного исследования с учетом данных на 21.04.2020 г., относительно эпидемии COVID-19 в г. Москва, можно сделать следующие выводы:

- 1) также как и в других регионах мира в Москве имеется значительное количество неучтенных пациентов, которые заражены COVID-19 и болеют в легкой или бессимптомной форме или уже выздоровели и имеют иммунитет к этой инфекции;
- 2) число умерших во время эпидемии превысит число жертв всех вместе взятых инфекционных заболеваний за то же период времени;
- 3) пик заболеваемости придется на период с 10.05.2020 г. по 05.06.2020 г.;
- 4) для достижения «фазового перехода» в эпидемическом процессе с резким сокращением числа заболевших умерших (китайский сценарий эпидемии) необходим комплекс мероприятий, направленный на массовую самоизоляцию населения и проведение масштабных мероприятий по дезинфекции.

Работа поддержана темой «АААА-А16-116122110060-9» бюджетного финансирования Лимнологического института СО РАН и темой «АААА-А17-117020210027-9» бюджетного финансирования Институт химической биологии и

фундаментальной медицины СО РАН. Авторы работы выражают благодарность Дмитрию Юрьевичу Щербакову за ценные советы, которые помогли выполнить работу.

Список литературы

- 1) Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., & Garry, R. F. (2020). The proximal origin of SARS-CoV-2. *Nature medicine*, 26(4), 450-452.
- 2) Velavan, T. P., & Meyer, C. G. (2020). The COVID-19 epidemic. *Tropical medicine & international health*, 25(3), 278.
- 3) He, R., Dobie, F., Ballantine, M., Leeson, A., Li, Y., Bastien, N., ... & Plummer, F. A. (2004). Analysis of multimerization of the SARS coronavirus nucleocapsid protein. *Biochemical and biophysical research communications*, 316(2), 476-483.
- 4) van Boheemen, S., de Graaf, M., Lauber, C., Bestebroer, T. M., Raj, V. S., Zaki, A. M., ... & Fouchier, R. A. (2012). Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. *MBio*, 3(6), e00473-12.
- 5) Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772), 700-721.
- 6) Bailey, N. T. (1975). *The mathematical theory of infectious diseases and its applications*. Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE.
- 7) Anderson, R. M., Anderson, B., & May, R. M. (1992). *Infectious diseases of humans: dynamics and control*. Oxford university press.
- 8) Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM review*, 42(4), 599-653.
- 9) Naresh, R., Tripathi, A., & Sharma, D. (2011). A nonlinear AIDS epidemic model with screening and time delay. *Applied Mathematics and Computation*, 217(9), 4416-4426.
- 10) Cai, L., Li, X., Ghosh, M., & Guo, B. (2009). Stability analysis of an HIV/AIDS epidemic model with treatment. *Journal of computational and applied mathematics*, 229(1), 313-323.
- 11) Liya, L. I. U., Jiang, D., Hayat, T., & Ahmad, B. (2018). Dynamics of a hepatitis B model with saturated incidence. *Acta Mathematica Scientia*, 38(6), 1731-1750.
- 12) Miller-Dickson, M. D., Meszaros, V. A., Almagro-Moreno, S., & Brandon Ogbunugafor, C. (2019). Hepatitis C virus modelled as an indirectly transmitted infection highlights the centrality of injection drug equipment in disease dynamics. *Journal of the Royal Society Interface*, 16(158), 20190334.
- 13) Reis, J., & Shaman, J. (2018). Simulation of four respiratory viruses and inference of epidemiological parameters. *Infectious Disease Modelling*, 3, 23-34.

- 14) Leecaster, M., Gesteland, P., Greene, T., Walton, N., Gundlapalli, A., Rolfs, R., ... & Samore, M. (2011). Modeling the variations in pediatric respiratory syncytial virus seasonal epidemics. *BMC infectious diseases*, 11(1), 105.
- 15) Osthus, D., Hickmann, K. S., Caragea, P. C., Higdon, D., & Del Valle, S. Y. (2017). Forecasting seasonal influenza with a state-space SIR model. *The annals of applied statistics*, 11(1), 202.
- 16) Gagnon, A., Miller, M. S., Hallman, S. A., Bourbeau, R., Herring, D. A., Earn, D. J., & Madrenas, J. (2013). Age-specific mortality during the 1918 influenza pandemic: unravelling the mystery of high young adult mortality. *PLoS One*, 8(8).
- 17) Baize, S., Pannetier, D., Oestereich, L., Rieger, T., Koivogui, L., Magassouba, N. F., ... & Tiffany, A. (2014). Emergence of Zaire Ebola virus disease in Guinea. *New England Journal of Medicine*, 371(15), 1418-1425.
- 18) Garske, T., Legrand, J., Donnelly, C. A., Ward, H., Cauchemez, S., Fraser, C., ... & Ghani, A. C. (2009). Assessing the severity of the novel influenza A/H1N1 pandemic. *Bmj*, 339, b2840.
- 19) Coburn, B. J., Wagner, B. G., & Blower, S. (2009). Modeling influenza epidemics and pandemics: insights into the future of swine flu (H1N1). *BMC medicine*, 7(1), 30.
- 20) Choi, B. C. K., & Pak, A. W. P. (2003). A simple approximate mathematical model to predict the number of severe acute respiratory syndrome cases and deaths. *Journal of Epidemiology & Community Health*, 57(10), 831-835.
- 21) Ng, T. W., Turinici, G., & Danchin, A. (2003). A double epidemic model for the SARS propagation. *BMC Infectious Diseases*, 3(1), 19.
- 22) Wu, P., Hao, X., Lau, E. H., Wong, J. Y., Leung, K. S., Wu, J. T., ... & Leung, G. M. (2020). Real-time tentative assessment of the epidemiological characteristics of novel coronavirus infections in Wuhan, China, as at 22 January 2020. *Eurosurveillance*, 25(3), 2000044.
- 23) Kucharski, A. J., Russell, T. W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., ... & Davies, N. (2020). Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The lancet infectious diseases*.
- 24) Puci, M. V., Loi, F., Ferraro, O. E., Cappai, S., Rolesu, S., & Montomoli, C. (2020). COVID-19 Trend Estimation in the Elderly Italian Region of Sardinia. *Frontiers in Public Health*, 8, 153.
- 25) Zhao, S., & Chen, H. (2020). Modeling the epidemic dynamics and control of COVID-19 outbreak in China. *Quantitative Biology*, 1-9.

- 26) Anastassopoulou, C., Russo, L., Tsakris, A., & Siettos, C. (2020). Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PloS one*, 15(3), e0230405.
- 27) Khrapov, P., & Loginova, A. (2020). Mathematical modelling of the dynamics of the Coronavirus COVID-19 epidemic development in China. *International Journal of Open Information Technologies*, 8(4), 13-16.
- 28) Rong, X., Yang, L., Chu, H., & Fan, M. (2020). Effect of delay in diagnosis on transmission of COVID-19. *Mathematical Biosciences and Engineering*, 17(3), 2725-2740.
- 29) DiCiccio, T. J., & Efron, B. (1996). Bootstrap confidence intervals. *Statistical science*, 189-212.
- 30) Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., ... & Xing, X. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England Journal of Medicine*.
- 31) Wu, J. T., Leung, K., Bushman, M., Kishore, N., Niehus, R., de Salazar, P. M., ... & Leung, G. M. (2020). Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nature Medicine*, 1-5.
- 32) Sanche, S., Lin, Y. T., Xu, C., Romero-Severson, E., Hengartner, N., & Ke, R. (2020). High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2. *Emerging infectious diseases*, 26(7).
- 33) Fraser, C., Donnelly, C. A., Cauchemez, S., Hanage, W. P., Van Kerkhove, M. D., Hollingsworth, T. D., ... & Jombart, T. (2009). Pandemic potential of a strain of influenza A (H1N1): early findings. *science*, 324(5934), 1557-1561.
- 34) Jüni, P., Rothenbühler, M., Bobos, P., Thorpe, K. E., da Costa, B. R., Fisman, D. N., ... & Gesink, D. (2020). Impact of climate and public health interventions on the COVID-19 pandemic: A prospective cohort study. *CMAJ*.

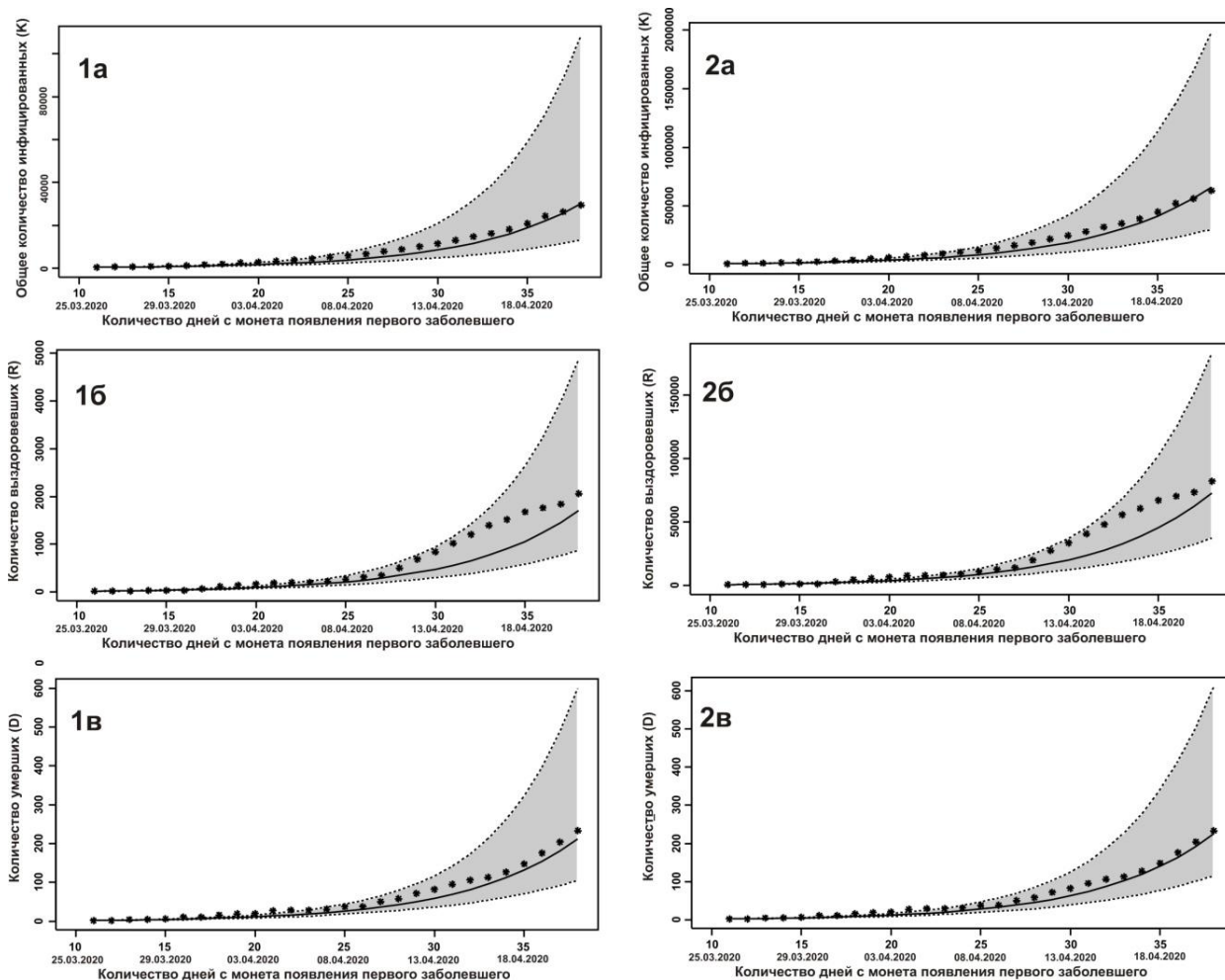


Рис. 1. Кривые, демонстрирующие поведение переменных $K(t)$, количества выздоровевших $R(t)$, количества умерших $D(t)$ и данных реальных наблюдений $K'(t)$, $R'(t)$, $D'(t)$ развития эпидемического процесса в г. Москва за период с 25.03.2020 г. по 21.04.2020 г. Сплошные линии показывают медианное значение $K(t)$, $R(t)$ и $D(t)$, пунктирные линии и серая область показывают 95% доверительные интервал. 1а, 1б и 1в – кривые первого сценария, построенные на наблюдаемых значениях $K'(t)$, $R'(t)$ и $D'(t)$. 2а, 2б и 2в – кривые второго сценария, построенные на скорректированных значениях, $K'(t)$ умножено на 20, $R'(t)$ умножено на 40.

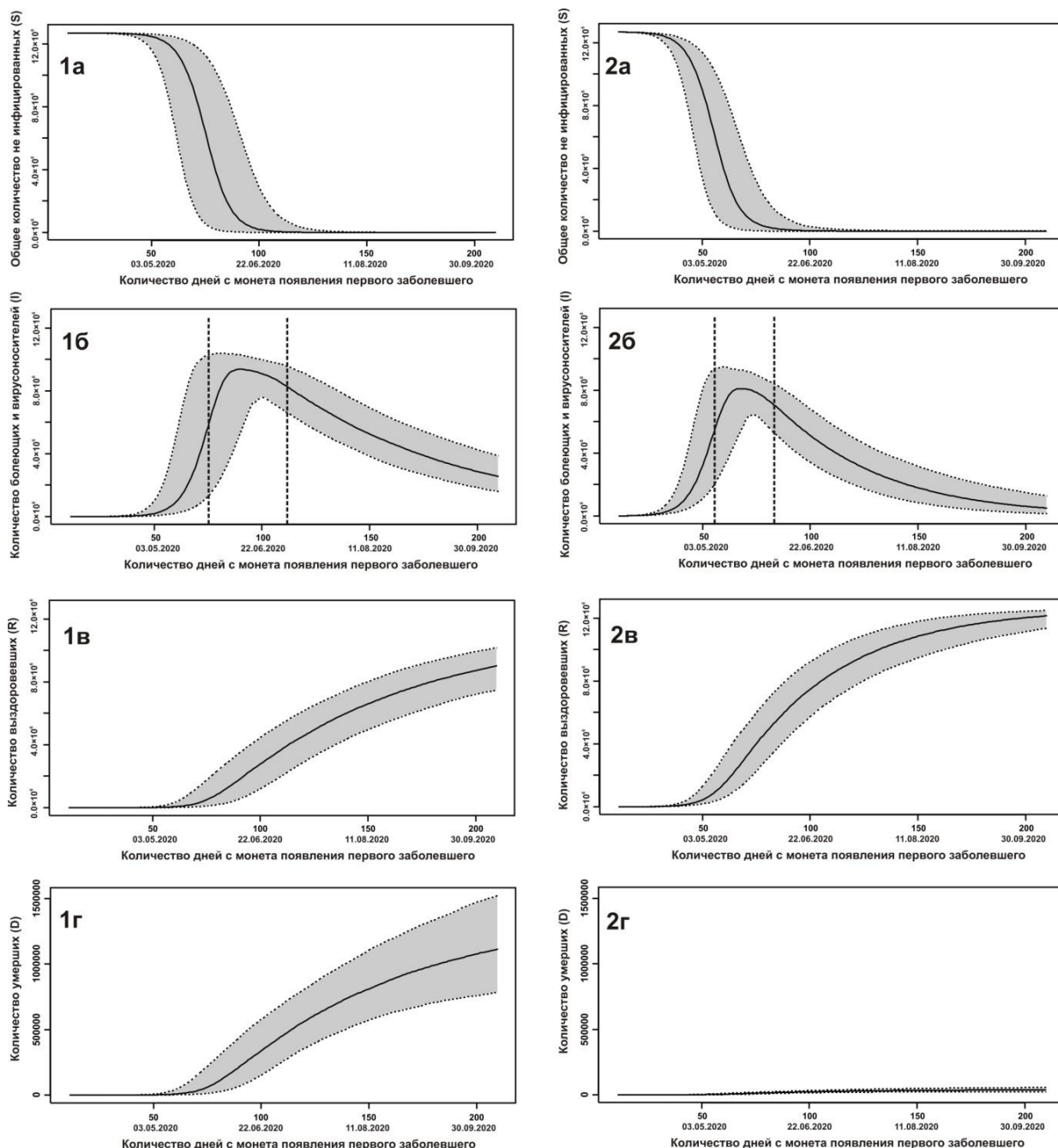


Рис. 2. Кривые, демонстрирующие предсказываемое поведение переменных количества человек из популяции, которые еще не проконтактировали с патогеном $S(t)$, количество болеющих или вирусоносителей $I(t)$, количество выздоровевших $R(t)$, количество умерших $D(t)$ в г. Москва за период 200 дней развития эпидемии с 15.03.2020 г. по 30.09.2020 г. Сплошные линии показывают медианное значение $S(t)$, $I(t)$, $R(t)$ и $D(t)$, пунктирные линии и серая область показывают 95% доверительный интервал. 1а, 1б, 1в и 1г – кривые первого сценария, рассчитанные на параметрах первого сценария без коррекции количества зараженных и выздоровевших. 2а, 2б, 2в и 2г – кривые второго сценария, построенные на параметрах второго сценария по скорректированным значениям количества зараженных и выздоровевших. Вертикальные пунктирные отрезки на 1в и 2в показывают доверительный интервал для пика заболеваемости.

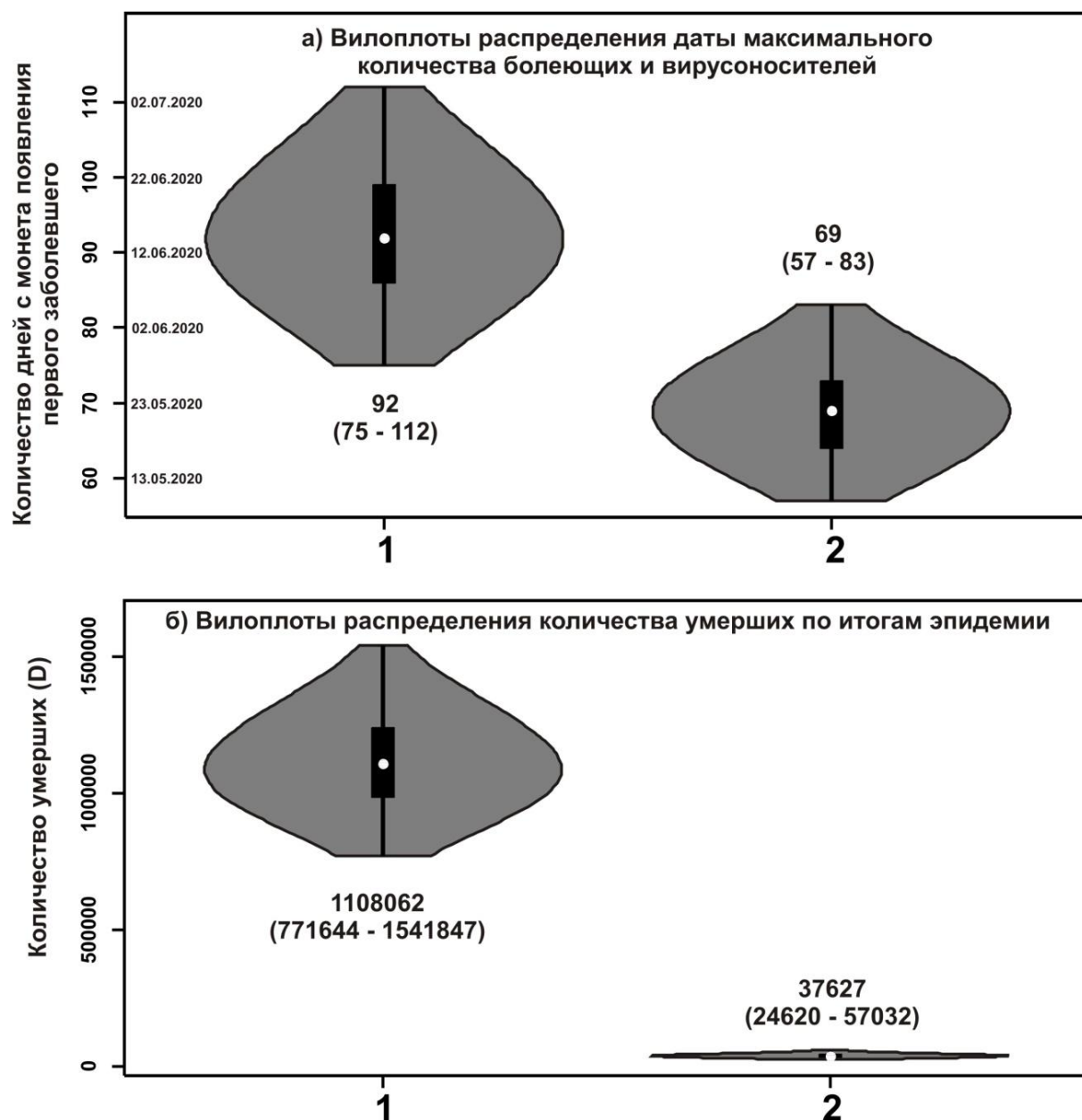


Рис. 3. Распределения наиболее вероятных моментов времени (даты) пика заболеваемости (количество болящих или вирусоносителей $I(t)$) и количества умерших после окончания эпидемии (значение переменной $D(t)$). Распределение (95% доверительный интервал) показанный в виде «виолплотов» - способ визуализации, дающий представление о форме распределения. 1 – данные полечены по сценарию без коррекции количества зараженных и выздоровевших. 2 - данные полечены по сценарию с коррекцией количества зараженных и выздоровевших.

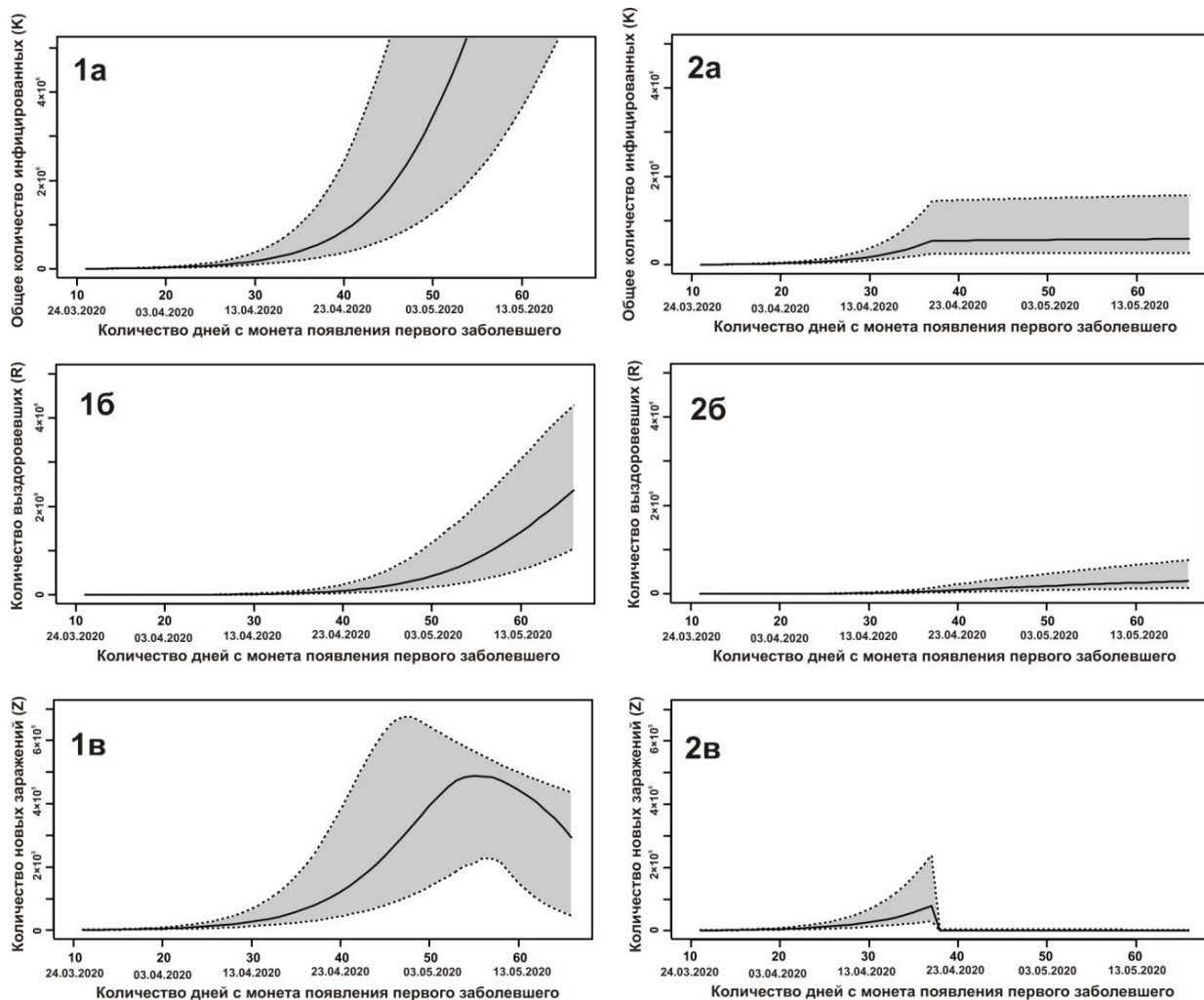


Рис. 4. Кривые, демонстрирующие предсказываемое поведение переменных количества инфицированных $K(t)$, количества выздоровевших $R(t)$, прироста числа заболевших за сутки $Z(t)$ в г. Москва в период с 15.03.2020 г. по 09.05.2020 г.. Сплошные линии показывают медианное значение $K(t)$, $R(t)$, и $Z(t)$, пунктирные линии и серая область показывают 95% доверительные интервал. 1а, 1б и 1в – кривые сценария, построенные на параметрах, оцененных по скорректированным значениям количества зараженных и выздоровевших (полная картина сценария приведена на рисунке 2 - 2а, 2б, 2в, 2г). 2а, 2б и 2в – кривые того же сценария с резким уменьшением значения параметра α в момент времени 21.04.2020 г. (α уменьшено в 55 раз). 2а, 2б и 2в соответствует сценарию с высокоэффективными противоэпидемическими мероприятиями.