

Мониторинг и прогноз динамики заболеваемости COVID-19 в Москве: 2020–2021 годы

А. В. Соколов^{*1}, Л. А. Соколова²

¹Институт проблем передачи информации им. А. А. Харкевича РАН, Москва

²Институт системного анализа Федерального исследовательского центра «Информатика и управление» РАН, Москва

Резюме

Актуальность. Накопление информации (статистических данных и знаний) о пандемии COVID-19 приводит к уточнению математических моделей, к расширению области их использования. **Целью** данного исследования является построение ряда моделей (соответствующих имеющимся знаниям и данным) для идентификации функций, определяющих динамику пандемии, и анализа возможностей проведения прогнозов. **Материалы и методы.** В работе использованы данные открытых статистических и информационных ресурсов, касающиеся всех аспектов COVID-19. Основой проведенного исследования является метод сбалансированной идентификации и одноименная информационная технология, созданная в Центре распределенных вычислений Института проблем передачи информации РАН. Технология используется для построения (выбора) моделей, соответствующих количеству и качеству данных, проведения расчетов (прогнозов) и представления результатов (в том числе все графики, приведенные ниже, подготовлены на ее основе). **Результат.** Построенные модели удовлетворительно описывают динамику заболеваемости COVID-19 в Москве. Они могут использоваться для прогноза с горизонтом в несколько месяцев при условии, что в моделируемом объекте не появляются новые, ранее отсутствовавшие элементы. Основным внутренним механизмом, определяющим динамику модели, является коллективный иммунитет и повышение заразности вируса (из-за распространения штаммов Дельта и Омикрон). **Заключение.** Приведены результаты успешного использования технологии сбалансированной идентификации для мониторинга пандемии COVID-19: модели, соответствующие данным, доступным в различные моменты времени (с марта 2020 по декабрь 2021); полученные новые знания – функциональные зависимости, определяющие динамику системы; расчеты различных эпидемических показателей (заболеваемость, иммунитет, индексы репродукции и др.); различные прогнозы для Москвы (от 01.12.2020, 15.04.2021, 01.08.2021 и 01.08.2021).

Ключевые слова: заболеваемость, COVID-19, моделирование, мониторинг, прогноз, сбалансированная идентификация
Конфликт интересов не заявлен.

Для цитирования: Соколов А. В., Соколова Л. А. Мониторинг и прогноз динамики заболеваемости COVID-19 в Москве: 2020–2021 годы. Эпидемиология и Вакцинопрофилактика. 2022;21(4): 48-59. <https://doi.org/10.31631/2073-3046-2022-21-4-48-59>

Monitoring and Forecasting the Dynamics of the Incidence of COVID-19 in Moscow: 2020–2021

AV Sokolov^{*1}, LA Sokolova²

¹Institute for information transmission problem (Kharkevitch Institute) RAS, Moscow

²Federal Research Center «Computer Science and Control» of RAS Institute for Systems Analysis, Moscow

Abstract

Relevance. The accumulation of information (statistical data and knowledge) about the COVID-19 pandemic leads to the refinement of mathematical models, to the expansion of the area of their use. **The aim** of this study is to build a set of models (in line with current knowledge and data) to identify the functions that drive the dynamics of a pandemic and analyze the possibilities for making predictions. **Materials and methods.** The work used data from open statistical and information resources relating to all aspects of COVID-19. The basis of the study is the balanced identification method and the information technology of the same name, created at the Center for Distributed Computing of the Institute for Information Transmission Problems of the Russian Academy of Sciences. The technology is used to build (select) models that correspond to the quantity and quality of data, perform calculations (forecasts) and present results (all the graphs below were prepared on its basis). **Result.** The constructed models satisfactorily describe the dynamics of the incidence of COVID-19 in Moscow. They can be used for a forecast with a horizon of several months, provided that new, previously absent elements do not appear in the modeled object. The main internal mechanism that determines the dynamics of the model is herd immunity and an increase in the infectivity of the virus (due to the spread of Delta and Omicron strains). **Conclusion.** The results of the successful use of balanced identification technology for monitoring the COVID-19 pandemic

* Для переписки: Соколов Александр Витальевич, к. ф.-м. н., ведущий научный сотрудник Института проблем передачи информации им. А. А. Харкевича РАН, 127051, Москва, Большой Каретный переулок, д. 19 стр. 1. +7-916-297-31-62, alexander.v.sokolov@gmail.com.
©Соколов А. В. и др.

** For correspondence: Sokolov Alexander V., Cand. Sci. (Phys.-Math.), Leading Researcher of Institute for information transmission problem (Kharkevitch Institute) RAS Bolshoy Karetny per. 19, build. 1, Moscow 127051 Russia. +7-916-297-31-62, alexander.v.sokolov@gmail.com.
©Sokolov AV, et al.

are presented: models corresponding to data available at various points in time (from March 2020 to December 2021); the acquired new knowledge - functional dependencies that determine the dynamics of the system; calculations of various epidemic indicators (morbidity, immunity, reproduction indices, etc.); various forecasts for Moscow (from 12/01/2020, 04/15/2021, 08/01/2021 and 08/01/2021).

Keywords: modeling, monitoring, forecast, dynamics, COVID-19, balanced identification

No conflict of interest to declare

For citation: Sokolov AV, Sokolova LA. Monitoring and Forecasting the Dynamics of the Incidence of COVID-19 in Moscow: 2020–2021. *Epidemiology and Vaccinal Prevention*. 2022;21(4): 48-59 (In Russ.). <https://doi:10.31631/2073-3046-2022-21-4-48-59>

Введение

Эффективное проведение мониторинга состояния и прогноза поведения сложных объектов подразумевает широкое использование математического моделирования. При определении целей мониторинга и прогноза необходимо учитывать наличие математических моделей, сложность, подробность и надежность которых, в свою очередь, определяются количеством и качеством (точностью) экспериментальных данных и знаний о функционировании объекта.

В случае пандемии COVID-19 по мере накопления данных и знаний и построения все более подробных и точных моделей, возможности мониторинга и прогноза возрастают. В доступных рядах данных проявляются различные эффекты, что позволяет усложнять соответствующие математические модели. Такому усложнению способствует и появление новых знаний об исследуемых процессах.

Динамика эпидемии определяется процессами взаимодействия вируса, человеческого организма и общества. Различные процессы имеют различные характерные времена, причем, чем больше характерное время процесса, тем больше времени требуется для его проявления, тем длиннее должны быть ряды наблюдений, необходимые для определения его характеристик.

Для пандемии COVID-19 характерные времена различных процессов могут быть оценены следующим образом:

- 15 дней – заразность и манифестация (проявляемость симптомов) в зависимости от длительности заболевания (времени, прошедшего с момента заражения);
- 30 дней – текущий индекс репродукции (контактное число) и индекс проявляемости симптомов (и последующей изоляции);
- 60 дней – зависимость индекса выявления от количества проведенных тестов;
- 90 дней – влияние естественного (после болезни) коллективного иммунитета;
- 200 дней – влияние ослабления естественного иммунитета со временем;
- 100 дней – влияние вакцинации;
- 180 дней (предварительная оценка) – влияние ослабления искусственного (после вакцинации) иммунитета,

- 365 дней – сезонность заразности,
- 60–150 дней – появление и распространение новых штаммов, изменение заразности.

Горизонт прогноза определяется погрешностью модели, которая определяется точностью описания отдельных процессов, которая, в свою очередь, определяется длиной рядов наблюдений (и их надежностью). Таким образом, длительность мониторинга определяет, какие процессы могут быть оценены (с характерными временами меньше этой длительности) и, следовательно, определяют горизонт и точность прогноза. С этой точки зрения становится понятной неудача попытки прогноза динамики пандемии, предпринятой Н. Фергюсоном. Его группа в начале 2020 г. активно обосновывала тотальный карантин, прогнозируя (на несколько месяцев вперед) 500 тыс. смертей в Англии и 2 млн в США (за лето 2020 г.) [<https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-03-16-COVID19-Report-9.pdf>, стр. 7].

На начальном этапе (первые месяцы пандемии), когда объем знаний и статистических данных был невелик и ненадежен, мы ограничились целью поиска закономерностей взаимодействия вируса и человека (заразность и манифестация вируса как функция длительности заболевания) и описания некоторых социальных механизмов управления эпидемией (количество контактов и изоляция зараженных). Рассматривались процессы с характерным временным периодом около 30 дней.

Затем, по мере накопления опыта (анализ траекторий модели и сравнение их с данными), знаний и статистической информации, был проведен ряд модификаций, позволивших учесть более медленные процессы (с большими характерными временами):

- эффективность выявления и изоляции зараженных в виде функции от количества проводимых (ежедневно) тестов (лето 2020 г.);
- естественный иммунитет переболевших (осень 2020 г.);
- потерю иммунитета в виде функции, зависящей от времени, прошедшего с момента заражения (зима 2020–2021);
- искусственный иммунитет (вакцинированных) и его потеря (весна–осень 2021).

Целью данного исследования является построение ряда моделей (соответствующих

Original Articles

имеющимся знаниям и данным) для идентификации функций, определяющих динамику пандемии, и анализа возможностей проведения прогнозов.

Материалы и методы

При построении моделей использовалась следующая статистическая информация:

- Новые случаи заражения в Москве – из официальной отчетности (см., например, <https://стопкоронавирус.пф/information/>).
- Количество проводимых тестов в России – из официальной отчетности (см., например, <https://tatspravka.ru/statistika-russia/>) Количество проведенных тестов на 1000 чел. в Москве оценивалось в 1/5 от проведенных в России.
- Процент носителей антител (имеющих иммунитет) в Москве – из заявлений официальных лиц и цитирований в интернете: <https://ria.ru/20200710/1574177822.html>, <https://www.vedomosti.ru/society/news/2020/09/28/841417-popova-nazvala-dolyu-moskvichei-s-antitelami-k-koronavirusu>, <https://www.vesti.ru/article/2533510>.
- Количество проведенных в Москве прививок – из официальной отчетности (см., например, <https://стопкоронавирус.пф/news/>).

Основой проведенного исследования стали метод сбалансированной идентификации [1] и одноименная информационная технология, созданная в Центре распределенных вычислений Института проблем передачи информации РАН. Технология используется для построения (выбора) моделей, соответствующих количеству и качеству данных, проведения расчетов (прогнозов) и представления результатов (в том числе все графики, приведенные ниже, подготовлены на ее основе). В настоящий момент доступна программная реализация технологии (<https://github.com/distcomp/SvF>).

Следует отметить, что основная переменная модели, полное количество инфицированных (выявленных и невыявленных), не является наблюдаемой (измеряемой) величиной. Наблюдаемая величина – новые выявленные случаи заражения (nC), которые являются (неизвестной) частью от полного количества. Это приводит к неоднозначности идентификации – внутренняя динамика модели определяется с точностью до множителя. Для преодоления этого недостатка в критерий сбалансированной идентификации добавляется специальное слабое, минимизирующее полное количество зараженных. Кроме того, для восстановления корректности постановки задачи (устранения неоднозначности идентификации) используются дополнительные данные – количество носителей антител, которое (по предположению модели) совпадает с полным количеством переболевших, сохранивших иммунитет.

В терминах принятых обозначений эпидемиологических моделей (см., напр. [2]) используются

различные модификации моделей SIR и $SIRS$, отличительной особенностью которых является разбиение инфицированных (I) на 15 групп, в соответствии с длительностью заболевания (Age of Infection Model, см. например, $SI_{15}R$ в Приложении). В терминах моделей популяционной динамики и демографии используются (матричные) модели Лесли [3] и модели Маккендрика фон Фёрстера для популяций с возрастной структурой [4,5].

Сложность моделей изменялась по мере появления новых данных и учета новых процессов (тестирование, вакцинация и т.п.). Специфические символично-числовые обозначения соответствующих модификаций отражают используемые данные и учитываемые процессы. Формальные описания моделей приведены в Приложении.

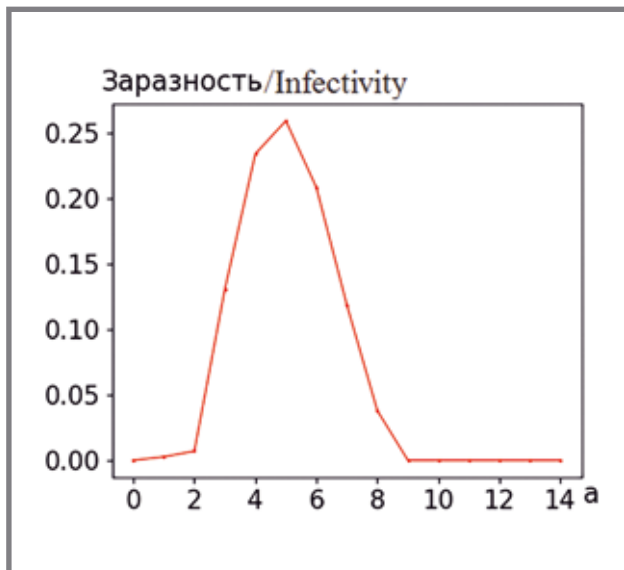
Эволюция моделей

В начале пандемии мы не знали о вирусе почти ничего, а доступная статистическая информация ограничивалась числами новых случаев заражения (nC) и смерти. На этом этапе мы ограничились целью выявления закономерностей взаимодействия вируса и человека – заразности и манифестации в зависимости от длительности болезни (времени, прошедшего с момента заражения). Для этого использовалась модель динамики зараженных $SI_{15}R-nC$ (см. Приложение), распределенных по длительности заражения. В качестве данных использовались только временные ряды выявленных зараженных (nC). Уже через пару месяцев на основе обработки временных рядов по 7 популяциям (Великобритания, Германия, Италия, Испания, Франция, Россия без Москвы и Московской области и Москвы с Московской областью) удалось оценить функции взаимодействия человека и вируса, определяющие динамику эпидемии, например, заразность как функцию от длительности (времени) болезни (рис. 1), а также оценить динамику социальных управлений – индекса репродукции RO_t (контактное число) и индекса выявления (и последующей изоляции) IS_t . Подробно с результатами можно ознакомиться в [6].

К лету 2020 г. накопился достаточный объем новой информации – число ежедневно проводимых тестов (nT). Это позволило поставить новую цель – описать усилия общества по выявлению (и последующей изоляции) больных в зависимости от количества проведенных тестов. Соответствующая модификация модели $SI_{15}R-nC-nT$ (см. Приложение) позволила оценить связь количества выявленных зараженных с количеством проведенных тестов. Найденная функция (с учетом бессимптомных и незарегистрированных носителей вируса) эффективности выявления в зависимости от количества тестов на тысячу человек для Москвы приведена на рисунке 2. Анализ кривой позволяет для данной модели оценить эффективность выявления больных в Москве величиной не более 17%, т.е. из полного числа инфицированных

Рисунок 1. Нормированная (интеграл равен единице) заразность как функция от длительности заражения (дни)

Figure 1. Normalized (the integral is equal to one) infectivity as a function of the duration of infection (days)



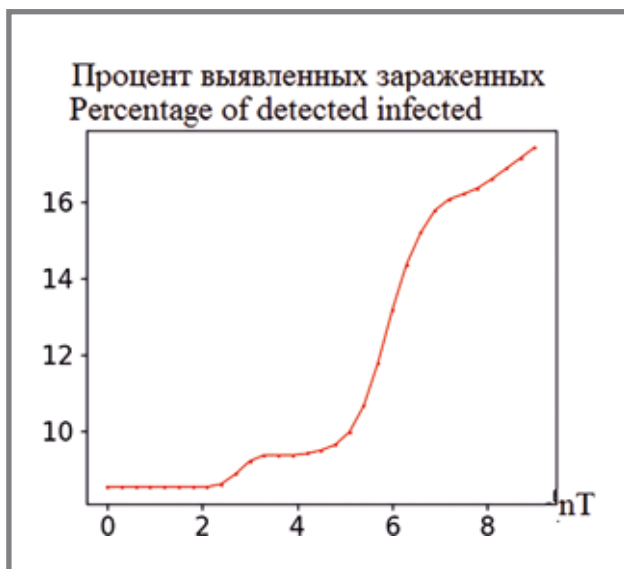
выявляется (и изолируется) не более одного заболевшего из шести.

Рассмотренные две модели были линейными, они определяли внутреннюю динамику с точностью до множителя и не учитывали размер популяции. Реальное количество инфицированных неизвестно (не является наблюдаемой величиной). Для оценки полного числа инфицированных нужна информация нового типа.

Летом 2020 г. такая информация появилась – (скрининговые) оценки числа носителей

Рисунок 2. Эффективность выявления больных как функция от количества тестов на тысячу человек (nT) в Москве

Figure 2. Patient detection efficiency as a function of the number of tests per thousand people (nT) in Moscow



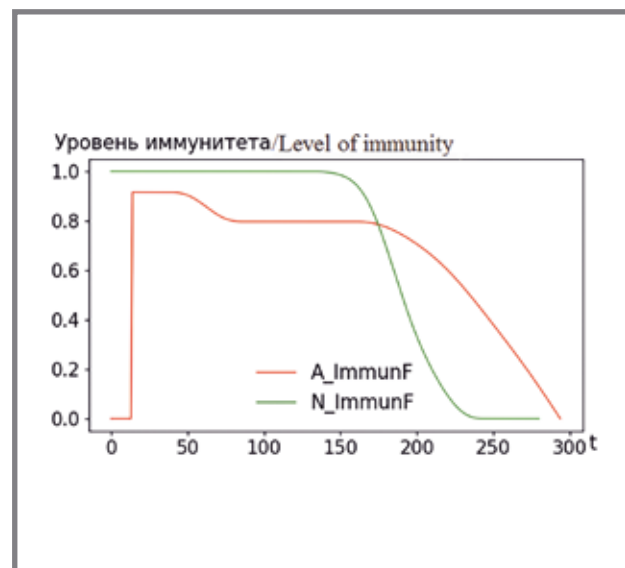
(значимого титра) антител, представляющие собой полное количество переболевших (а не только выявленных) за время пандемии (точнее за время, когда сохраняются антитела). Соответствующая модификация модели $SI_{15}R-nC-nT-Anti$ позволила для Москвы оценить реальное количество зараженных (не только выявленных), учесть количество переболевших (считается, что у них появляются антитела и естественный иммунитет), что существенно улучшило прогностические возможности модели. Модель перестала быть линейной, появилось саморегулирование процесса.

Однако в декабре 2020 г. пространство решений модели стало недостаточно емким для описания наблюдаемых данных, что вызвало очередную ее модификацию – учет потери естественного (полученного в результате заболевания) иммунитета. Идентификация новой модификации модели $SI_{15}RS-nC-nT-Anti-Im$ позволила определить уровень естественного иммунитета как функцию времени от начала заболевания (рис. 3) и оценить среднее время его утраты в 190 дней.

Наконец, с 2021 г. началась массовая вакцинация. Для ее учета в модификации модели $SI_{15}RS-nC-nT-Anti-Im-Vuc$ добавлен множитель, уменьшающий поток инфицированных в зависимости от процента населения, имеющего искусственный (в результате вакцинации) иммунитет. Идентификация модели позволила определить уровень искусственного иммунитета как функцию времени от вакцинации (см. рис. 3). При этом предполагалось, что первые 14 дней иммунитет от вакцинации отсутствует. Среднее время утраты иммунитета от вакцинации оценивается в 200 дней.

Рисунок 3. Уровень иммунитета как функция времени (дни)

Figure 3. The level of immunity as a function of time (days)



Примечание: N_ImmunF – естественный иммунитет, A_ImmunF – искусственный иммунитет.

Note: N_ImmunF – natural immunity, A_ImmunF – artificial immunity.

Original Articles

В этой модификации для управления эпидемией имеется уже три средства (медицинского и социального) управления: вакцинация, тестирование и ограничение контактов.

Результаты и обсуждение

К декабрю 2020 г. накопился достаточный опыт мониторинга пандемии, и основные процессы, определяющие её динамику (на тот момент), были выявлены, описаны и идентифицированы. Построенные в результате модели использовались для прогнозирования.

На рисунках 4–10 приведены результаты моделирования, идентификации, прогноза и верификации новых (выявленных) случаев. Начало моделирования 19.03.2020. На оси t нанесены отметки, соответствующие месяцам и годам. Вертикальная желтая линия соответствует дате начала прогноза. Модельные кривые (красного цвета) были получены (идентифицированы) на обучающем наборе данных (синие точки). Построенная таким образом модель использовалась для прогноза. При этом управление модели (индекс репликации, количество новых тестов и количество новых вакцинированных), как правило, замораживалось на последнем (перед началом прогноза) значении. Для верификации модели использовался полученный позже и незадействованный при построении модели тестирующий набор данных (зеленые точки).

Прогноз от 01.12.2020 – конец второй волны

Согласно модели, падение числа выявленных случаев во вторую волну пандемии

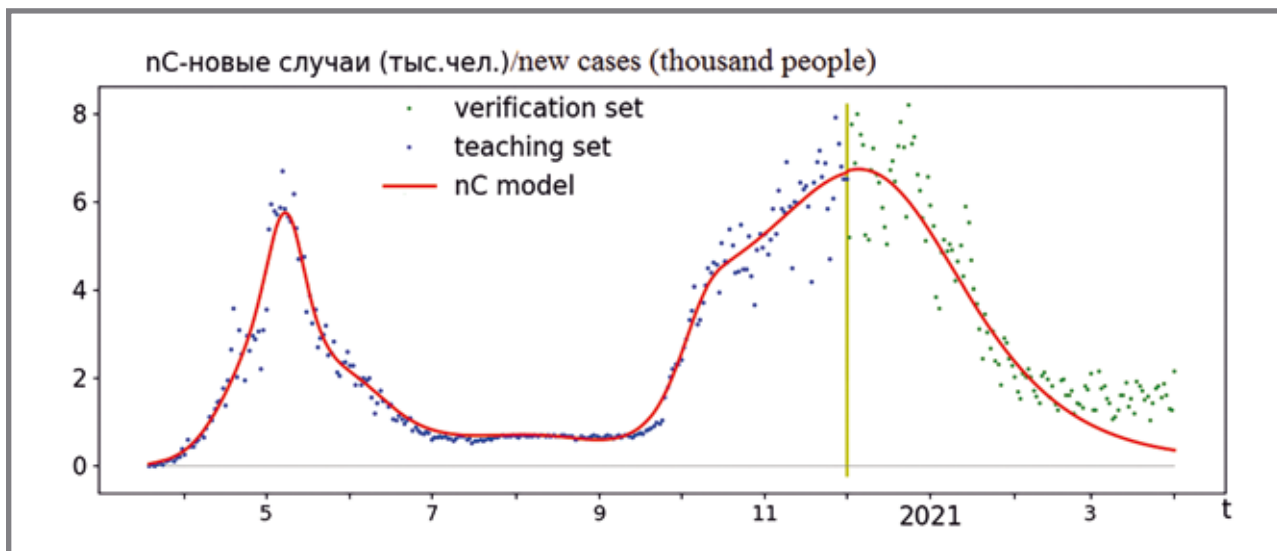
(декабрь–январь 2021 г.) в большой степени определяется накоплением иммунной прослойки (населения, имеющего естественный иммунитет). Наблюдаемое с середины февраля расхождение (рис. 4) можно объяснить существенным ослаблением (противоэпидемических) ограничений и/или миграцией заболевших из других регионов России.

Прогноз от 15.04.2021 – третья волна

На рисунке 5 приведен прогноз третьей волны в Москве при заданных управлениях пандемической ситуацией: контакты и тесты замораживаются на значениях 2.43 и 4.48 соответственно (от 15 апреля 2021 г.), вакцинация – 12 000 чел./день (эффективность вакцинации 91,6%). Предполагается, что иммунитет от вакцинации со временем не изменяется, т.е. коллективный искусственный иммунитет определяется количеством привитых (рис. 6).

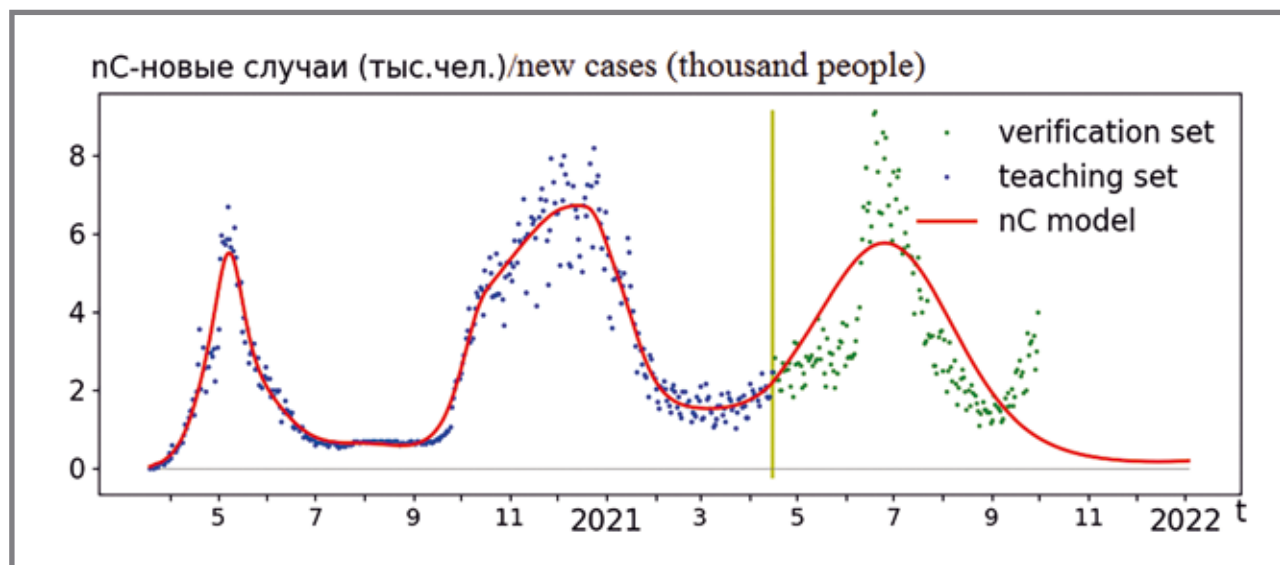
В течение последующих пяти месяцев прогноз качественно совпадает с реальными данными, которые не использовались при идентификации модели (зеленые точки). Согласно модели, третья волна (увеличение числа выявленных случаев с мая 2021 г.) в большой степени определяется уменьшением иммунной прослойки переболевшего населения. Действительно, с мая – начинают терять иммунитет переболевшие в начале второй волны (октябрь–ноябрь). В июне рост выявленных зараженных резко усиливается, что соответствует потере естественного иммунитета значительным количеством переболевших в декабре 2020 г. Периодичность волн составляет около 200 дней и в основном определяется функцией сохранения

Рисунок 4. Прогноз новых случаев заражения в Москве от 01.12.2020
Figure 4. Forecast of new cases of infection in Moscow from 12/01/2020



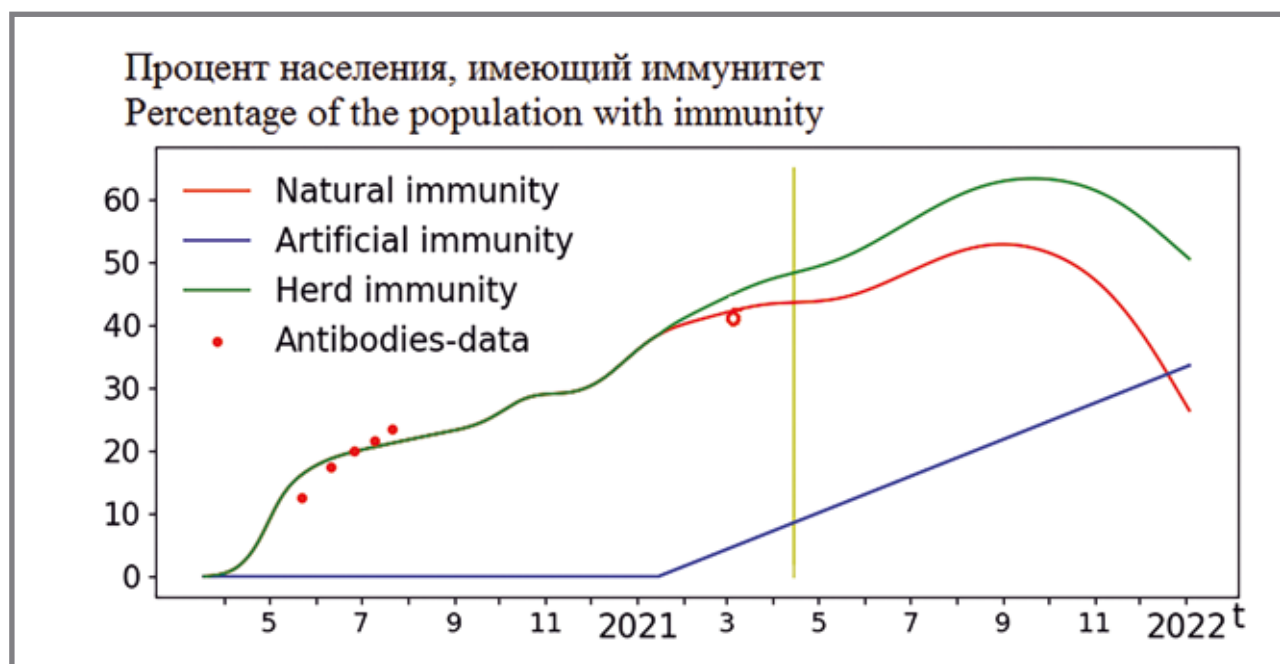
Примечание: Начало моделирования 19.03.2020. На оси t нанесены отметки, соответствующие месяцам и годам. Вертикальная желтая линия соответствует дате начала прогноза. Модельные кривые (красного цвета), полученные (идентифицированные) на обучающем наборе данных (синие точки) можно сравнить с тестирующим набором (зеленые точки).
Note: Start of simulation 19/03/2020. The t -axis has marks corresponding to months and years. The vertical yellow line corresponds to the start date of the forecast. The model curves (red) obtained (identified) on the training data set (blue dots) can be compared with the testing set (green dots).

Рисунок 5. Прогноз новых случаев заражения во время третьей волны эпидемии в Москве от 15.04.2021
Figure 5. Forecast of new cases of infection in the third wave of the epidemic in Moscow from 04/15/2021



Примечание: как к рисунку 4.
 Note: as to figure 4.

Рисунок 6. Иммунизация населения Москвы
Figure 6. Immunization of the population of Moscow



Примечание: Herd immunity – коллективный иммунитет; Artificial immunity – вакцинированные; Natural immunity – естественный иммунитет; Antibodies-data – данные по проценту носителей антител: точка, отмеченная красным кругом, использовалась для верификации.
 Note: Antibodies-data - data on the percentage of antibody carriers: the point marked with a red circle was used for verification.

иммунитета (см. рис. 3). Интенсивные мероприятия по временному уменьшению индекса репликации и увеличению тестирования (и последующей изоляции) могут несколько сдвинуть волну или изменить ее форму. При этом суммарное количество инфицированных меняется незначительно – основным механизмом, определяющим динамику, является общий (переболевших и вакцинированных) коллективный иммунитет. Согласно модели

($SI_{15}RS-nC-nT-Anti-Im-Vuc$), для того чтобы волна пошла на спад, коллективный иммунитет должен превысить некоторое значение, которое определяется интенсивностью тестирования и индексом репродукции (ограничительными мерами общества). Для первой волны это значение составляло около 15% населения, для второй – около 35%, для третьей оценивается значением около 55%. Такое увеличение можно объяснить

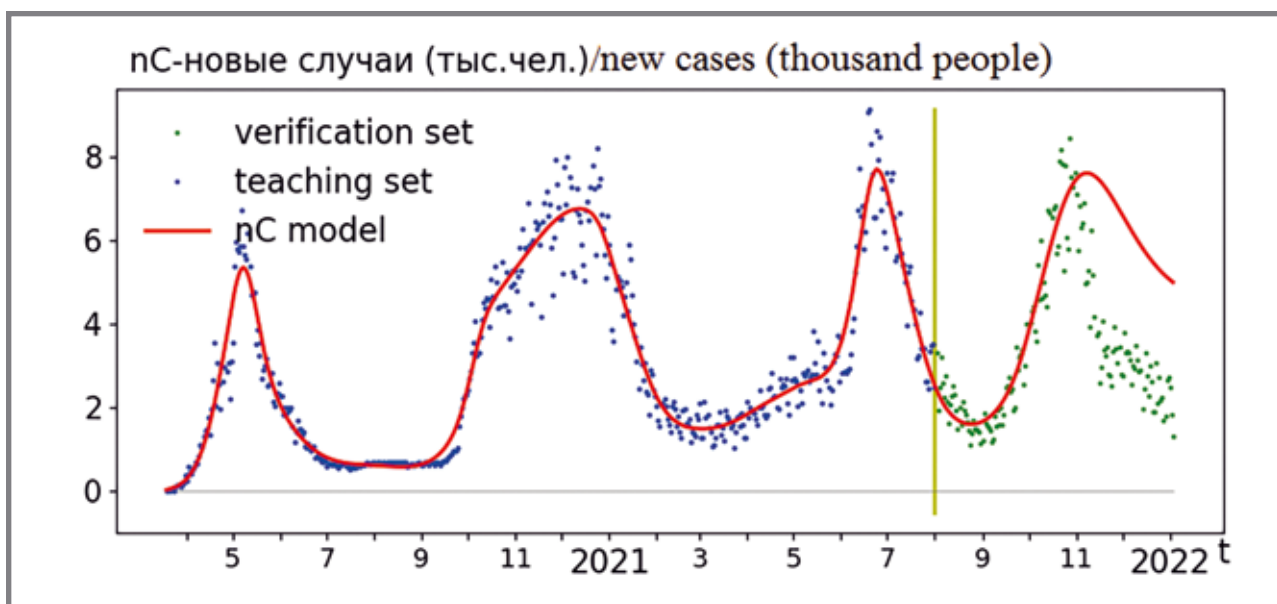
Original Articles

изменением (мутацией) вируса в направлении большей заразности.

При вакцинации 12 тыс. в день прогноз количества выявленных зараженных в третьей волне (с 15.04.2021 по 01.11.2021) составляет 671 тыс. человек. Соответствующая оценка смертности (при условии, что умирают 2,5% от выявленных инфицированных) составляет около 16 тыс. Введение (временных) ограничений, уменьшающих

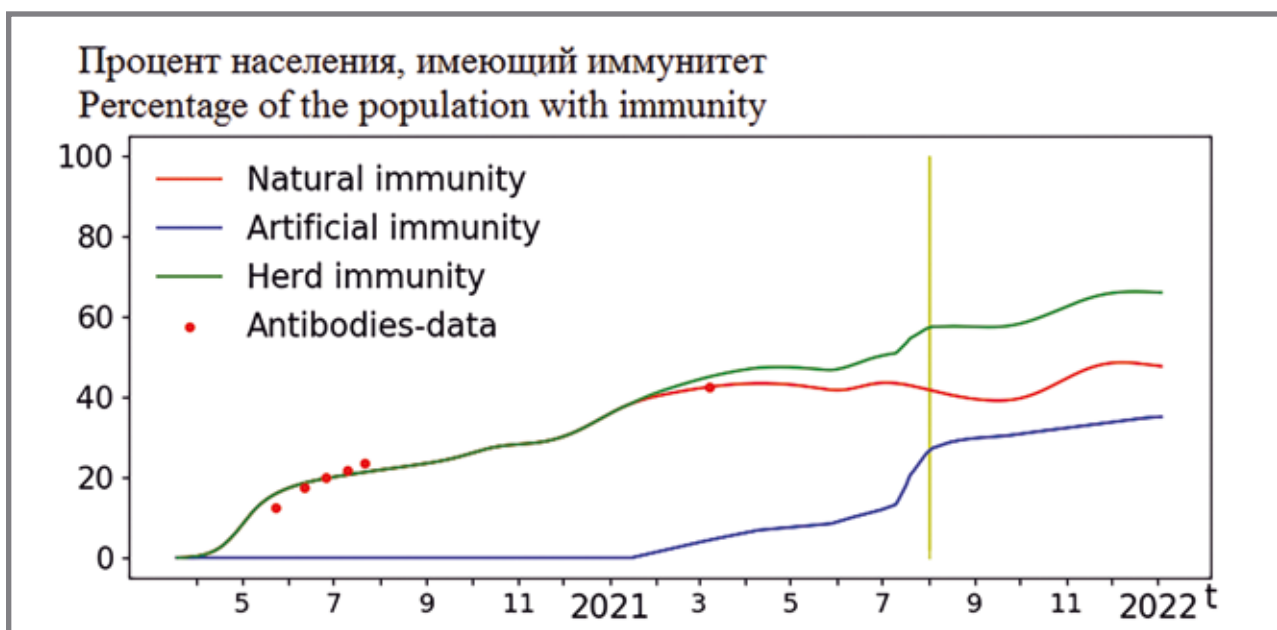
RO_t , и увеличение тестирования растягивает волну и уменьшают ее высоту, однако прогнозируемое суммарное количество выявленных зараженных в третьей волне меняется несущественно. Это некая константа, которую можно существенно уменьшить лишь вакцинацией – по модельным расчетам (оценка на 17.06.2021), 11 вакцинированных уменьшают количество выявленных переболевших на одного человека. Таким образом,

Рисунок 7. Прогноз новых случаев заражения в четвертой волне эпидемии в Москве от 01.08.2021
Figure 7. Forecast of new cases of infection in the fourth wave of the epidemic in Moscow from 08/01/2021



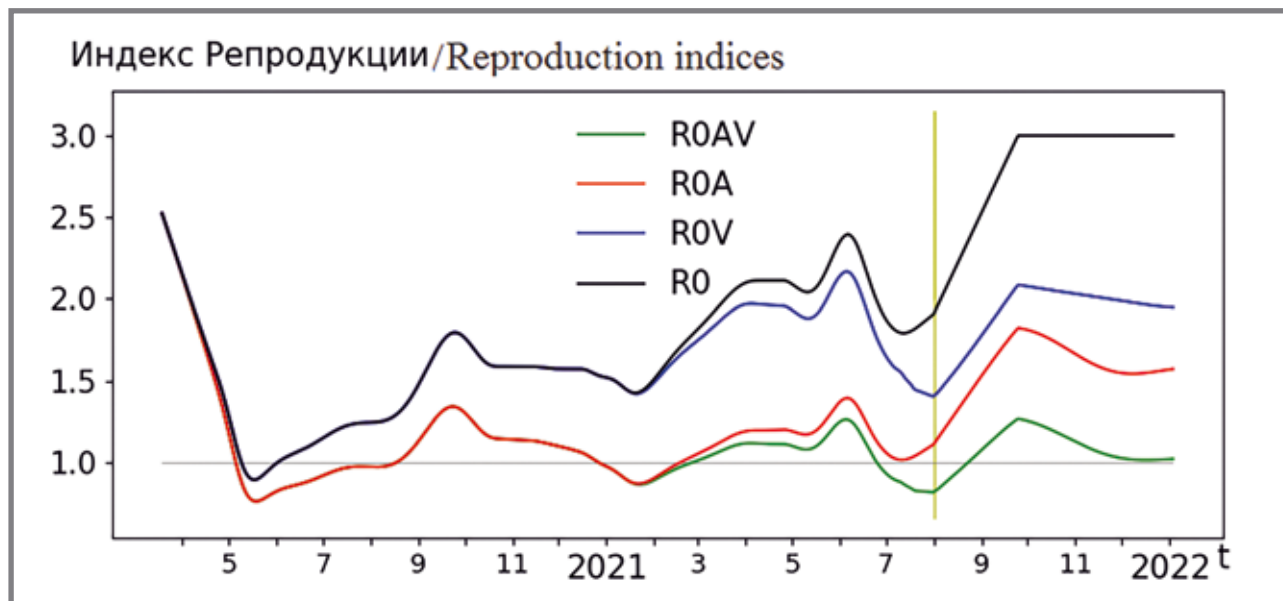
Примечание: как к рисунку 4.
 Note: as to figure 4.

Рисунок 8. Иммунизация населения Москвы. Прогноз от 01.08.2021
Figure 8. Immunization of the population of Moscow. Forecast from 08/01/2021



Примечание: как к рисунку 6.
 Note: as to figure 6.

Рисунок 9. Индексы репродукции R_0 (прогноз от 01.08.2021)
Figure 9. Reproduction indices R_0 (forecast from 08/01/2021)



Примечание: R_0A – индекс репродукции с учетом иммунитета переболевших, R_0V – индекс репродукции с учетом вакцинации, R_0AV – индекс репродукции с учетом иммунитета переболевших и вакцинации.

Note: R_0A is the reproduction index taking into account the immunity of those who have been ill, R_0V is the reproduction index taking into account vaccination, R_0AV is the reproduction index taking into account the immunity of those who have been ill and vaccination.

вакцинация дополнительных (по сравнению с прогнозом из расчета 12 тыс. в день) 1,5 млн человек уменьшает количество выявленных заболевших на 135 тыс. человек. При смертности 2,5% получаем, что 440 привитых уменьшают количество умерших в третью волну эпидемии на одного человека, соответственно 1,5 млн дополнительных вакцинированных спасают около 3400 жизней.

Согласно официальной статистике, количество зараженных, выявленных в Москве с 15.04.2021 по 19.07.2021, составило 410 тыс. человек, что неплохо согласуется с прогнозом в 671 тыс. человек, особенно если принять во внимание, что непредусмотренное прогнозом интенсивное вакцинирование уменьшило количество зараженных на 120 тыс. чел. (оценка).

На рисунке 6 приведены другие результаты моделирования – динамика коллективного иммунитета и его составляющих: процент вакцинированных (соответствует темпам вакцинации 12 тыс. в день) и прогноз процента переболевших, сохранивших к моменту времени естественный иммунитет. Точка, отмеченная красным кругом (оценка процента переболевших, имеющих иммунитет на 08.03.2021), не использовалась при идентификации. Близость её к траектории можно рассматривать как дополнительное подтверждение пригодности модели для прогноза (верификацию модели). Полученная кривая неплохо согласуется с результатами более подробного исследования динамики выявляемости IgG, приведенными в [7].

Прогноз от 01.08.2021 – четвертая волна

На рисунках 7–9 представлен более реалистичный прогноз, построенный на основе

модификации модели $SI_{15}RS-nC-nT-Anti-Im-Vuc$, которая учитывает ослабление искусственного иммунитета со временем (в соответствии с функцией см. на рис. 3). Прогноз предполагает ослабление ограничений при сохранении старых темпов вакцинирования и тестирования:

- интенсивность вакцинации 18 тыс. чел./день,
- индекс репродукции (R_0 на рис. 9) поднимается к началу сентября к значению 3,0, которое соответствует отсутствию ограничений (для штамма Дельта),
- количество тестов к ноябрю падает до умеренного значения 4,5 в день на тысячу населения,
- предполагается небольшой приток инфицированных извне в Москву. Их полное число составляет 600 чел. в день, что увеличивает количество ежедневно выявляемых инфицированных примерно на 90 человек.

На рисунке 7 приведен прогноз новых случаев заражения и реальные данные. В течение пары месяцев наблюдается неплохое совпадение, однако далее возникает заметное расхождение. Его можно объяснить существенным уменьшением индекса репродукции (непредусмотренного в прогнозе) из-за принятых мер по ограничению контактов и ревакцинации.

На рисунке 8 приведен прогноз уровня коллективного иммунитета и его составляющих для населения Москвы в четвертую волну эпидемии COVID-19 (от 01.08.2021). Прогноз предполагает довольно высокий темп вакцинации – 18 тыс. в день.

Однако с учетом эффекта потери иммунитета от вакцинации (см. рис. 3) искусственный иммунитет населения растет недостаточно быстро.

Original Articles

Аналогично, потеря естественного иммунитета переболевших в третью волну значительно уменьшает долю населения, имеющего естественный иммунитет в сентябре и октябре. В результате коллективный иммунитет не увеличивается, что с ростом заразности (штамм Дельта) приводит к росту заболеваемости и к четвертой волне эпидемии.

Рассмотренные составляющие иммунизации, заразность вируса и меры по ограничению передачи вируса определяют различные составляющие индекса репродукции (рис. 9).

Заразность вируса и ограничительные меры определяют R_0 – (первичный) индекс репродукции, который характеризует интенсивность контактов. Другие индексы получаются из R_0 путем учета процента иммунизированного населения: ROA – индекс репродукции с учетом иммунитета переболевших, ROV – индекс репродукции с учетом вакцинации, $ROAV$ – индекс репродукции с учетом иммунитета переболевших и вакцинации. Именно значение $ROAV$ является основным фактором определяющим динамику заболеваемости.

Прогноз от 01.12.2021 – Омикрон

На рисунке 10 приведен прогноз от 01.12.2021 количества выявленных заболевших в Москве при заданных управлениях: индекс репликации (контакты) и тесты замораживаются на значениях 2,25 и 6,28 соответственно, вакцинация – 20 тыс. в день. Прогноз хорошо совпадает с реальными данными только в течение одного месяца до вспышки, вызванной резким увеличением заразности (штамма Омикрон), не предусмотренной в модели.

Заключение

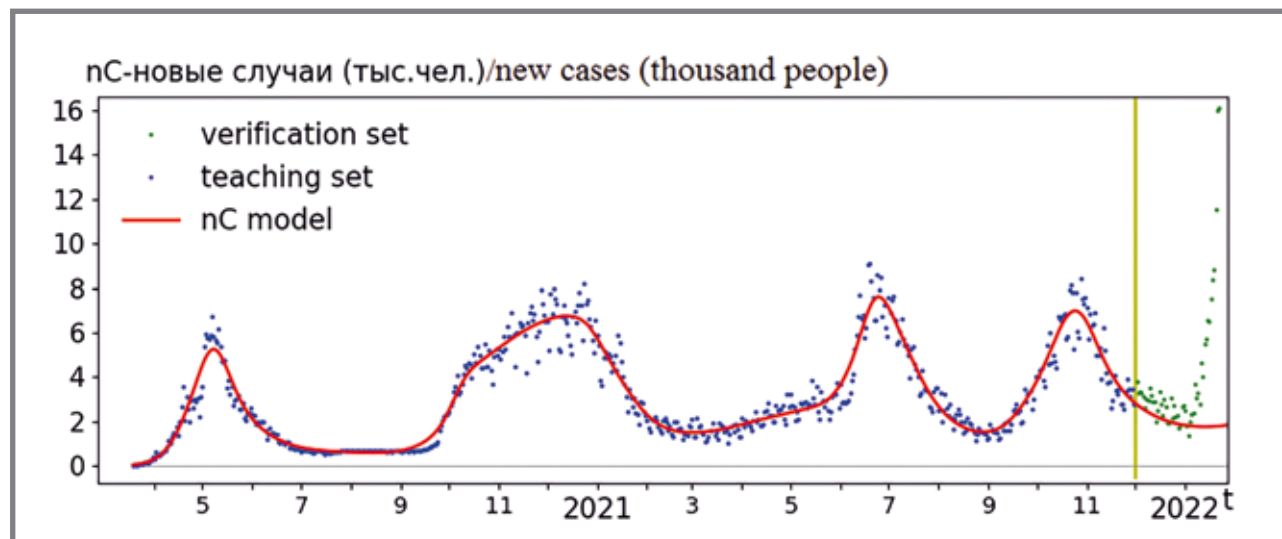
По мере наблюдения за пандемией COVID-19 в исследуемых рядах данных проявлялись

различные эффекты. Их учет позволил последовательно (шаг за шагом) усложнять модель. В результате удалось построить более или менее полную (на 2020–2021 гг.) модель, удовлетворительно описывающую динамику эпидемии. Сравнение полученных прогнозов с реальными данными показывает неплохое совпадение в течение нескольких месяцев. Проявившиеся при этом погрешности, как правило, объясняются неопределенностью сценариев (будущих) медико-социальных управлений, необходимых для расчетов (ограничение, выявление и вакцинация) или проявлением новых эффектов, не учтенных в модели. Таким образом, модель может использоваться для прогнозов с горизонтом несколько месяцев.

Процесс модификации модели на сегодняшний день не закончен. Трудно предсказать появление новых лекарств, вакцин и штаммов, однако нужно быть готовыми оперативно вносить в модель соответствующие изменения. Предполагается, что следующая модификация модели будет описывать смену штаммов.

Все приведенные результаты (модели, расчеты, прогнозы, включая приведенные здесь графики) получены на основе использования технологии сбалансированной идентификации. Используемая технология позволяет организовать эволюционный процесс построения моделей – от простого к сложному. Такая постепенная (эволюционная) модификация является крайне желательной – рассматриваемые постановки (задачи идентификации) являются сложными двухуровневыми оптимизационными задачами (возможно многоэкстремальными) и их решение требует значительных ресурсов. Поэтому поиск решения без правдоподобного начального приближения потребовал бы слишком больших вычислительных ресурсов, и, кроме того, нет оснований

Рисунок 10. Прогноз выявленных случаев заражения в Москве от 01.12.2021
Figure 10. Forecast of detected cases of infection in Moscow from 12/01/2021



Примечание: как к рисунку 4.
Note: as to figure 4.

предполагать, что найденное в этом случае решение (один из локальных минимумов оптимизационной задачи) будет иметь предметную интерпретацию, удовлетворяющую исследователей.

Такое последовательное усложнение задачи позволяет значительно экономить вычислительные ресурсы. Большая часть результатов данной работы рассчитана на современном ноутбуке (процессор CORE i5) за приемлемое время (до 15 минут). Для решения более сложных, ресурсоемких задач

использовались высокопроизводительные многопроцессорные комплексы [8].

Исследование выполнено с использованием оборудования центра коллективного пользования «Комплекс моделирования и обработки данных исследовательских установок мега-класса» НИЦ «Курчатовский институт», <http://ckp.nrcki.ru/>,

Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов 20-57-82004 и 20-07-00701

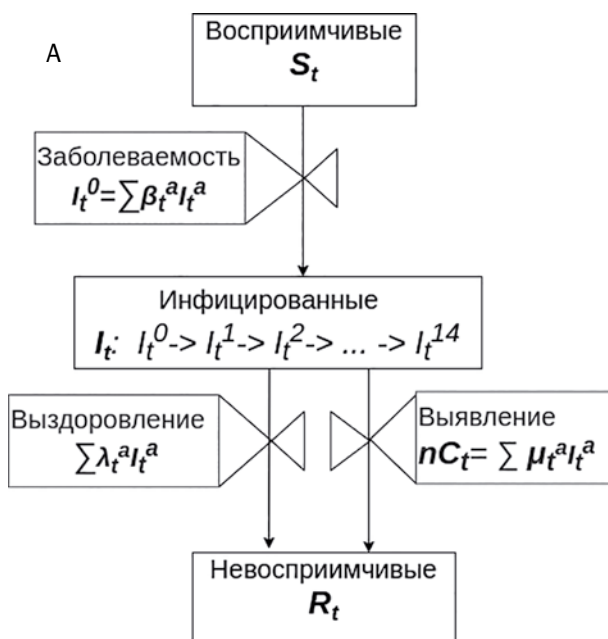
Литература

1. Sokolov A.V., Voloshinov V.V. Model Selection by Balanced Identification: the Interplay of Optimization and Distributed Computing. *Open Computer Science*, 2020, 10, p. 283–295. doi:10.1515/comp-2020-0116
2. Brauer F., Castillo-Chavez C., Feng Z. *Mathematical Models in Epidemiology*. Springer New York, NY; 2019. <https://doi.org/10.1007/978-1-4939-9828-9>
3. Свиричев Ю. М., Логофет Д. О. Устойчивость биологических сообществ. М.: Наука; 1978.
4. Нахушев, А. М. Уравнения математической биологии. М.: Высшая Школа; 1995.
5. Эбелинг В., Энгель А., Файстель Р. Физика процессов эволюции. М.: Эдиториал УРСС; 2001.
6. Sokolov A.V., Sokolova L.A. COVID-19 dynamic model: balanced identification of general biological and country specific features. // 9th International Young Scientist Conference on Computational Science (YSC 2020), *Procedia Computer Science*, Volume 178, 2020, Pages 301–310. doi:10.1016/j.procs.2020.11.032
7. Годков М. А., Шустов В. В., Коршунов В. А. и др. Формирование коллективного иммунитета к SARS-CoV-2 в популяции населения Москвы. *Эпидемиология и Вакцинопрофилактика*. 2022;21(1):81–91. <https://doi.org/10.31631/2073-3046-2022-21-1-81-91>
8. Sukhoroslov O., Volkov S., Afanasiev A. A Web-Based Platform for Publication and Distributed Execution of Computing Applications. *Parallel and Distributed Computing*, 14th International Symposium on IEEE. 2015. С. 175–184.

References

1. Sokolov AV, Voloshinov VV. Model Selection by Balanced Identification: the Interplay of Optimization and Distributed Computing. *Open Computer Science*, 2020;10:283–295. doi: 10.1515/comp-2020-0116
2. Brauer F, Castillo-Chavez C, Feng Z. *Mathematical Models in Epidemiology*. Springer New York, NY. 2019. <https://doi.org/10.1007/978-1-4939-9828-9>
3. Svirezhev YM, Logofet DO. *Sustainability of biological communities*. Moscow: Nauka; 1978 (In Russ).
4. Nakhushev AM. *The equations of mathematical biology. Textbook manual for universities*. Moscow: Higher School; 1995 (In Russ).
5. Ebeling V, Engel A, Feistel R. *Physics of evolutionary processes*. Moscow: Editorial URSS; 2001 (In Russ).
6. Sokolov AV, Sokolova LA. COVID-19 dynamic model: balanced identification of general biological and country specific features. 9th International Young Scientist Conference on Computational Science (YSC 2020), *Procedia Computer Science*. 2020;178:301–310. doi:10.1016/j.procs.2020.11.032
7. Godkov MA, Shustov VV, Korshunov VA, et al. Formation of Herd Immunity to SARS-CoV-2 in the Population of Moscow. *Epidemiology and Vaccinal Prevention*. 2022;21(1):81–91 (In Russ.). <https://doi.org/10.31631/2073-3046-2022-21-1-81-91>
8. Sukhoroslov O, Volkov S, Afanasiev A. A Web-Based Platform for Publication and Distributed Execution of Computing Applications. *Parallel and Distributed Computing*, 14th International Symposium on IEEE. 2015:175–184.

Приложение. Математические модели



SI15R. Разностная модель. Согласно рисунку A1, в данной модели население разбивается на 3 основные группы: S – восприимчивые, I – инфицированные (выявленные+невыявленные) и R – невосприимчивые (выздоровевшие и изолированные). Кроме того, группа инфицированных разбивается на 15 подгрупп в соответствии со временем, прошедшим с момента заражения a (age of infection).

Формализация рассмотренной схемы приводит к разностной модели:

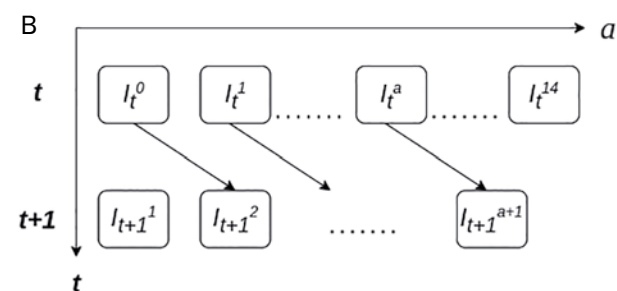


Рисунок A1. А. Модель SI15R с разбиением инфицированных на 15 групп. В. Иллюстрация динамики подгрупп инфицированных, которая определяется сдвигом по возрасту заболевания (a)

Figure. A1. A. Model SI15R with division of the infected into 15 groups. B. Illustration of the dynamics of subgroups of infected, which is determined by the shift in the age of the disease (a).

Original Articles

$$I_{t+1}^{a+1} = I_t^a - \lambda^a \cdot I_t^a - \mu_t^a \cdot I_t^a, \quad a \in [0, 14]; t \in [0, T]$$

$$I_{t+1}^0 = B_t; t \in [0, T]$$

$$B_t = \sum_{a=0}^{14} \beta_t^a \cdot I_t^a; t \in [0, T]$$

$$I_0^a = I_0^a; a \in [0, 14]$$

$$nC_t = \sum_{a=0}^{14} \mu_t^a \cdot I_t^a; t \in [0, T]$$

где a – длительность (возраст) заболевания, t – время моделирования, – количество (невыявленных) зараженных в момент времени t с длительностью заражения a , – вероятность выздороветь, – вероятность выявления и последующей изоляции, B_t – количество новых зараженных, – заразность инфицированного, – распределение инфицированных по времени заражения в начальный момент времени. – новые (выявленные) случаи заражения.

Непрерывная модель.

$$\frac{\partial I}{\partial t} + \frac{\partial I}{\partial a} = -\lambda(a)I(t, a) - \mu(t, a)I(t, a)$$

$$I(t, 0) = \int_0^{a_{\max}} \beta(t, a') I(t, a') da';$$

$$I(0, a) = I_0(a)$$

$SI_{15}R-nC$. Основная идея этой модификации модели $SI_{15}R$ состоит в мультипликативном представлении функций выявления и заразности:

$$\mu_t^a = IS_t \cdot m^a; \sum_{a=0}^{14} m^a = 1; \beta_t^a = RO_t \cdot b^a; \sum_{a=0}^{14} b^a = 1$$

где – доля выявленных и изолированных от общего числа зараженных, – зависимость выявления от длительности заражения, – текущее репродуктивное контактное число (сколько в среднем заражает один инфицированный, при условии, что он не будет выявлен и изолирован), – заразность от длительности заражения (a).

Дополнительно, предполагаются достаточно очевидные ограничения на функции :

$$\lambda^0=0; \lambda^1=0; m^0=0; m^1=0; b^0=0; b^1=0; b^{14}=0$$

$$IS_{t+1} \geq IS_t; t \in [0, T].$$

Последнее неравенство (не убывание доли выявленных инфицированных) отражает увеличение усилий общества по выявлению и изоляции больных (в первые месяцы эпидемии).

$SI_{15}R-nC-nT$. Данная модификация модели связывает усилия общества по выявлению (и последующей изоляции) больных в зависимости от количества проведенных тестов.

$$IS_t = IS_{nT} (nT_t)$$

$SI_{15}R-nC-nT-Anti$. Учет естественного иммунитета переболевших (естественного коллективного иммунитета).

Количество новых зараженных

$$B_t = (1 - N_Immunity_t) \cdot RO_t \cdot \sum_{a=0}^{14} b^a \cdot I_t^a; t \in [0, T]$$

$$N_Immunity_t = 1/P_{mos} \sum_{t'=0}^t B(t'),$$

где – доля носителей антител в момент времени t , P_{mos} – население Москвы.

$SI_{15}RS-nC-nT-Anti-Im$. Учет уменьшения естественного коллективного иммунитета. В отличие от предыдущей версии, при расчете естественного иммунитета учитывается уменьшение иммунитета со временем.

Доля населения, имеющего естественный иммунитет

$$N_Immunity_t = 1/P_{mos} \sum_{t'=0}^t N_Immunity_{t-t'} \cdot B_{t-t'},$$

где – функция сохранения естественного иммунитета как функция времени, прошедшего с момента заболевания (Δt): 1 – полный иммунитет (вероятность заболеть равна 0), 0 – полное отсутствие иммунитета.

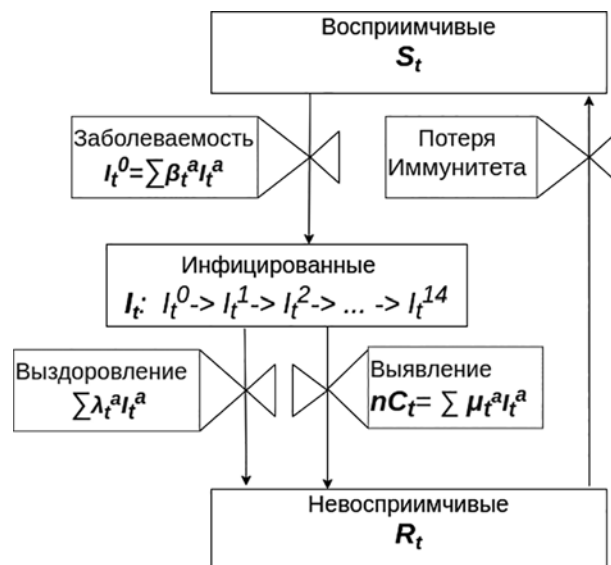


Рисунок A2. Модель SI15RS. Невосприимчивые (имеющие иммунитет) снова становятся восприимчивыми из-за потери иммунитета.

Figure. A2. Model SI15RS. Unsusceptible (those who have immunity) become susceptible again due to the loss of immunity.

$SI_{15}RS-nC-nT-Anti-Im-Vuc$. Учет вакцинации. Учет искусственного иммунитета осуществляется аналогично учету естественного иммунитета в виде дополнительного множителя, уменьшающего заболеваемость

$$B_t = (1 - A_Immunity_t) \cdot (1 - N_Immunity_t) \cdot R_0 \cdot \sum_{a=0}^{14} b_{a,t}^a \cdot I_t^a; t \in [0, T],$$

где $VUCC_t$ – интенсивность вакцинации (количество прививок в день) в момент t , – функция сохранения искусственного иммунитета как функция времени, прошедшего с момента вакцинации (Δt): 1 – полный иммунитет (вероятность заболеть равна 0), 0 – полное отсутствие иммунитета.

Об авторах

- **Александр Витальевич Соколов** – к. ф.-м. н., ведущий научный сотрудник Института проблем передачи информации им. А. А. Харкевича РАН, 127051, Москва, Большой Каретный переулок, д.19 стр. 1. +7-916-297-31-62, alexander.v.sokolov@gmail.com. ORCID: 0000-0003-0107-2519
- **Любовь Александровна Соколова** – техник Института системного анализа Федерального исследовательского центра «Информатика и управление» РАН. + 7-985-981-14-84, las.sokolova@gmail.com.

Поступила: 11.05.2022. Принята к печати: 07.07.2022.

Контент доступен под лицензией CC BY 4.0.

About the Authors

- **Alexander V. Sokolov** – Cand. Sci. (Phys.-Math.), Leading Researcher of Institute for information transmission problem (Kharkevitch Institute) RAS Bolshoy Karetny per. 19, build.1, Moscow 127051 Russia. +7-916-297-31-62, alexander.v.sokolov@gmail.com. ORCID: 0000-0003-0107-2519.
- **Lubov A. Sokolova** – technician of Federal Research Center «Computer Science and Control» of RAS Institute for Systems Analysis. + 7-985-981-14-84, las.sokolova@gmail.com.

Received: 11.05.2022. Accepted: 07.07.2022.

Creative Commons Attribution CC BY 4.0.