



OPEN

## Delay-differential SEIR modeling for improved modelling of infection dynamics

I. N. Kiselev<sup>1,2,3</sup>✉, I. R. Akberdin<sup>2,3,4</sup> & F. A. Kolpakov<sup>1,2,3</sup>

SEIR (Susceptible–Exposed–Infected–Recovered) approach is a classic modeling method that is frequently used to study infectious diseases. However, in the vast majority of such models transitions from one population group to another are described using the mass-action law. That causes inability to reproduce observable dynamics of an infection such as the incubation period or progression of the disease's symptoms. In this paper, we propose a new approach to simulate the epidemic dynamics based on a system of differential equations with time delays and instant transitions to approximate durations of transition processes more correctly and make model parameters more clear. The suggested approach can be applied not only to Covid-19 but also to the study of other infectious diseases. We utilized it in the development of the delay-based model of the COVID-19 pandemic in Germany and France. The model takes into account testing of different population groups, symptoms progression from mild to critical, vaccination, duration of protective immunity and new virus strains. The stringency index was used as a generalized characteristic of the non-pharmaceutical government interventions in corresponding countries to contain the virus spread. The parameter identifiability analysis demonstrated that the presented modeling approach enables to significantly reduce the number of parameters and make them more identifiable. Both models are publicly available.

Mathematical modeling of the spread of infectious diseases is a powerful and widely used approach to predict infection, lethality and mortality rates in a certain country or over the world<sup>1</sup>. It may also help reveal what should be the most effective administrative strategies and social containment measures in order to minimize loss of life and productivity and to curb the spread<sup>2–5</sup>. The SIR (Susceptible, Infected and Recovered) method is an often used approach to build epidemiological models<sup>6</sup>. It has been widely applied to simulate the epidemic spread, control mechanisms and impact on economic output of the COVID-19 in different countries<sup>7–10</sup>. SIR-models can be easily extended, for example, to include different aspects of the disease. For instance, in<sup>11</sup> inclusion of the viral load and the impact on the immune human system into the SIR-model has enabled the identification of potential causes of two-phase exponential growth of the epidemic. Another natural extension is taking into account incubation time of the virus. Such models are usually called SEIR-models where E means exposed<sup>12</sup>.

SEIR-models found enormous application in theoretical studies of diverse aspects of the novel SARS-CoV-2 pandemic. In particular, such models were used to estimate the impact of different lockdown intensities on epidemic spread in China<sup>13,14</sup>, United Kingdom<sup>15</sup> and Europe (e.g. the Netherlands)<sup>16</sup> for the year 2020, and even until 2025 for the USA, considering seasonal forcing and cross-immunity from the other betacoronaviruses<sup>17</sup>. The SEIR modeling has also been harnessed to estimate the effect of local and international travel restrictions on the spread of COVID-19 outbreak<sup>18</sup>.

The typical scenario of authorities' actions (also called NPIs for Non-Pharmaceutical Interventions) simulated in SEIR models is a restriction on the mobility and mass gatherings which reduce the number of contacts in the population and can be represented as an additional multiplier to the infection rate law reflecting social distancing<sup>16</sup>. The numerical value of the control parameter may be changed via discrete events and piecewise functions. Other key NPIs are closing borders and quarantine on entry which diminishes the influx of infected individuals to the simulated region and mass testing for the virus where different modes of testing can be implemented in the model depending on the financial capabilities and government acts and policies (random tests or testing of infected with severe/critical symptoms, or considering contacts of an infected individual etc.). The number of hospital beds and intensive care units (ICU) is another crucial factor in the fight of authority against COVID-19 which should be considered in epidemiological modeling<sup>19</sup>.

<sup>1</sup>FRC for Information and Computational Technologies, Novosibirsk, Russia. <sup>2</sup>Sirius University of Science and Technology, Sirius, Russia. <sup>3</sup>BIOSOFT.RU, Ltd, Novosibirsk, Russia. <sup>4</sup>Novosibirsk State University, Novosibirsk, Russia. ✉email: [axec@systemsbiology.ru](mailto:axec@systemsbiology.ru)

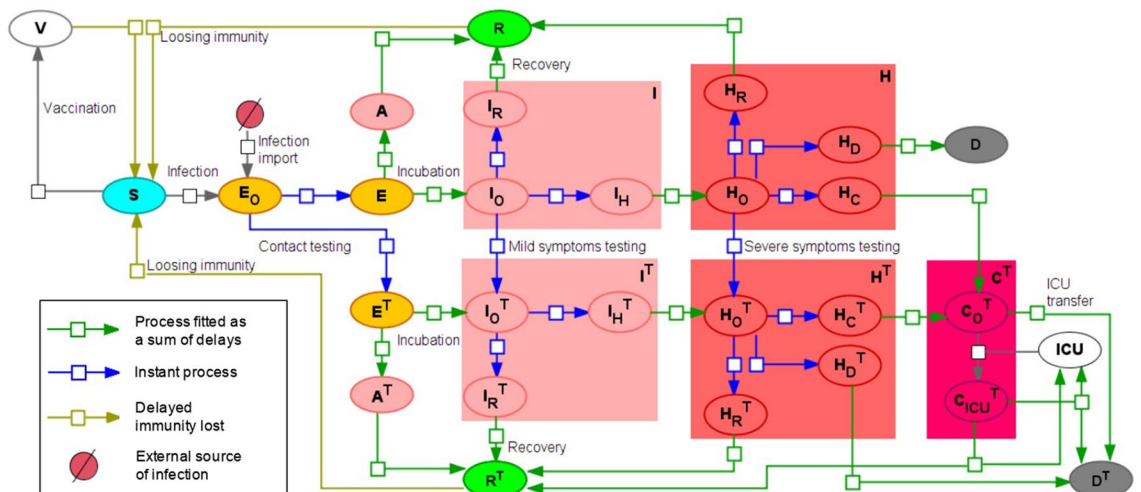
Despite the fact that initial results of the numerical study of SEIR models played an essential role in determining both basic laws of the primary development of the COVID-19 pandemic and core characteristics of the current pandemic situation, in the overwhelming majority this type of models use mass action laws to describe the transitions between states (for example, from the incubation period to the symptomatic). Because of that, such models cannot always adequately reproduce the dynamics of such transitions. The methodological constraint of the SEIR models can be solved by using delayed differential equations which are able to explicitly capture the durations of the latent, quarantine, and recovery periods<sup>20,21</sup>. Thus, Shayak and coauthors numerically investigated the simplest retarded logistic equation with time delay to model the spread of COVID-19 in a city and demonstrated that solution of the model is significantly sensitive to small changes in the parameter values<sup>22</sup>. At the same time, more conventional SEIR-based delay differential equation models were proposed to reproduce the COVID-19 dynamics in Germany, China, South Korea, India and Japan<sup>23–26</sup> and to predict the epidemic dynamics in Italy and Spain when it was in its early stages. However, these models did not take into account asymptomatic carriers and non-testing subpopulations as well as the progression of the disease's severity.

Herein, we propose novel mathematical model based on the model developed by<sup>16</sup> using differential equations with weighted sums of delayed argument mixed with instant processes, which allow us not only model transition processes adequately to clinically observed data, but also directly quantify the proportion of hospitalized patients with moderate and severe symptoms, on an ICU, asymptomatic, tested and untested among them, which can be compared with the available statistics. The main goal of the study is to present a new approach to model epidemiological processes combining delay-differential terms and instant processes which may be fitted separately from the rest of the model. This approach reduces the number of model parameters and makes them more epidemiologically interpretable and more identifiable compared to classic SEIR approach. The results of numerical analysis and model validation are demonstrated by the example of two European countries, Germany and France.

### Results

**Model structure.** The final version of the proposed delay differential equations (DDE) model consists of the following subpopulations or groups (Fig. 1):

1. *S*—susceptible to the SARS-CoV-2 virus.
2. *V*—vaccinated subpopulation, considered to be immune to the virus.
3. *E*—exposed to the virus. After the incubation period they will transit either to asymptomatic or symptomatic. Here we do not use additional subgroups due to equal time intervals for both transitions.
4. *A*—asymptomatic individuals, which will recover over time but can infect others<sup>27</sup>.
5. *I*—mild symptomatic group. It comprises three subgroups: with onset symptoms (*I<sub>O</sub>*) then they are instantly divided into those who will recover (*I<sub>R</sub>*) and those who will progress to the severe symptomatic (*I<sub>H</sub>*). Transition is done according to the fraction of severe symptomatic among those who show any symptoms (*F<sub>H</sub>*).
6. *H*—severe symptomatic group which comprises four subgroups: 1) with just onset symptoms (*H<sub>O</sub>*). They instantaneously transit into subgroups of individuals who will eventually recover (*H<sub>R</sub>*), die (*H<sub>D</sub>*) or progress to critically ill (*H<sub>C</sub>*). Transitions are performed according to the disease lethality (*F<sub>D</sub>*) and the fraction of critically ill (*F<sub>C</sub>*).
7. *C*—critically ill group where ICU is required in order to recover. If no ICU is available these patients will die. All critically ill patients are considered to be automatically tested for the virus infection.
8. *R*—recovered from COVID-19.



**Figure 1.** Overall SEIR-like model with instant and delayed processes. All abbreviations of the population groups described in the main text.

9.  $D$ —deceased due to COVID-19.
10. All infected subgroups (except critically ill) also have “registered” or “tested for COVID-19” counterparts:  $A^T, E^T, I^T, H^T, R^T, D^T$ . In the model patients may be tested at three different stages (1) when being exposed to the virus. It is done through contact tracing procedures. Percentage of exposed to the virus who will be tested and registered is set by  $T_E$  parameter; (2) upon symptoms onset. Percentage of mildly symptomatic individuals who will be registered is given by parameter  $T_I$ ; (3) upon severe symptoms onset. Percentage of severely symptomatic individuals who will be registered is given by parameter  $T_H$ .

Most transitions in the model are described as either instant processes or as preliminary fitted processes (blue and green arrows, correspondingly, in Fig. 1). To fit delayed processes we used data from<sup>28</sup> for incubation period and<sup>29</sup> for other epidemiological processes.

Process of release from a hospital may be fitted using data provided by Our World in Data for France for recovery\ dying in hospitals. To this end we constructed a partial model describing the process of admitting hospital, transition to ICU and leaving hospital (Fig. 2). Given daily numbers of the hospital admission, daily number of ICU admissions and daily number of hospital patients we fitted the process of leaving hospital utilizing formula (3) (see “Methods”). It should be noted that we assume that all severely ill patients are tested and moved to the hospital. However, it is not always the case and should be addressed in the updated version of the model.

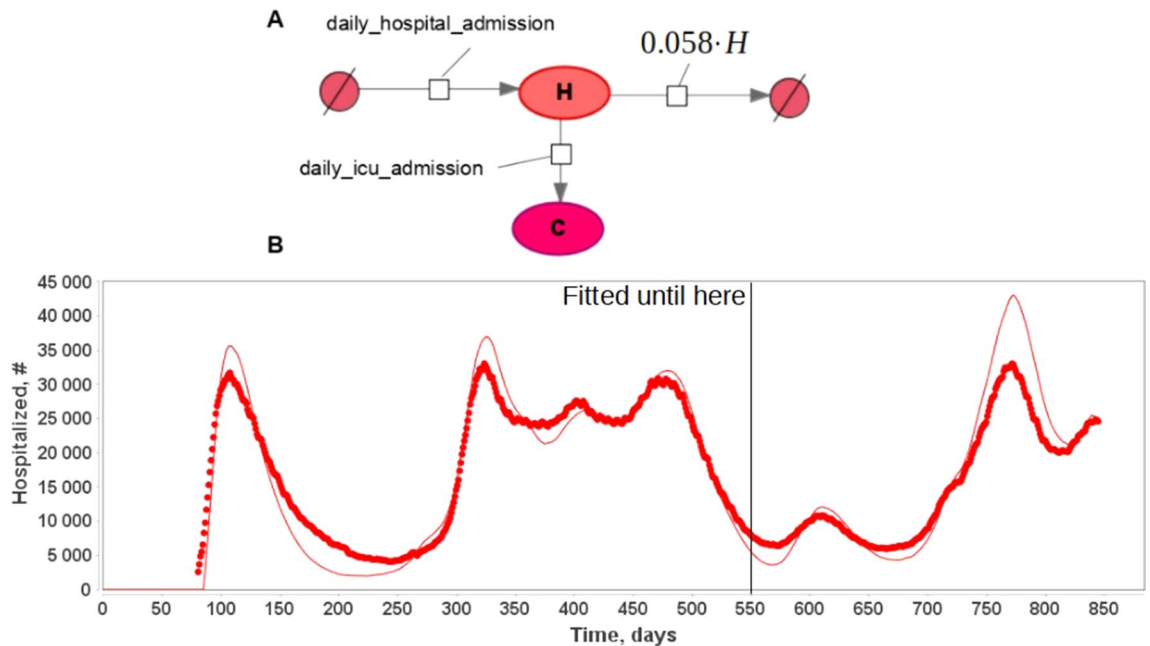
Overall scheme of the hospitalization model in SBGN format as well as a result of the model fitting to statistical data on hospitalization in France are presented in Fig. 2. It should be noted that hospital stay duration is shown to be the same for the whole pandemic duration and not dependent on the virus strain.

There are also four transitions in the model treated differently:

1. *Duration of the protective immunity.* We imply that the immunity acquired either after recovery or vaccination lasts 180 days according to the average experimental evaluations<sup>30,31</sup>.
2. *Vaccination.* Based on the known statistical data we established the number of individuals vaccinated each day in a certain country. For model purposes we allow the vaccination only for susceptible individuals (either never infected before or those who lost their immunity through time). The kinetic law for this process is zeroth-order:  $dV/dt = k_V$  with  $k_V$  changing each day based on the tabular data for the certain country. We also assume that vaccines have 100% efficiency immediately after the first dose, and the immunity via vaccination declines according to the duration period (see item 1).
3. *ICU admittance.* We considered this process to require a free ICU and be instant in most cases. However, lower value of the kinetic constant may be used to reflect the fact that not everyone who needs ICU gets it:

$$dC_{ICU}/dt = k_{ICU} \cdot ICU \cdot C, k_{ICU} \leq K_{Large}.$$

4. *The infection process.* Transition from susceptible to exposed is defined using Total Infection Coefficient (TIC) which we calculated similar to the model described in<sup>16</sup> with modifications:



**Figure 2.** (A): Partial model of the hospitalization due to Covid-19 using delay equations for hospital release. Data on daily hospital and ICU admission taken from ourworldindata.org. (B): Results of the model fitting to the number of hospitalized patients in France. Model was fitted until time = 550 days.

$$dE/dt = TIC \cdot S \cdot \frac{1}{SIF},$$

$$TIC = \frac{CF \cdot IC}{P} \cdot \sum_{j \in \{A, E, I, H\}} (j + Q_j \cdot j^T),$$

where  $CF$ —average number of contacts for individuals per day in the simulated region which can be obtained from statistical surveys,  $IC$ —a probability to be infected upon contact,  $Q_j$ —quarantine coefficient for corresponding group. Only registered individuals are subject to quarantine,  $SIF$  is a stringency index factor calculated based on the Stringency Index<sup>32</sup> which reflects government NPIs and imposed limits such as mask regime, limit on mass gatherings, school closing etc. Stringency index ranges from 0 (no interventions) to 100 (maximum possible interventions). In the current study we recalculated it to SIF as follows:

$$SIF = 100 / (100 - SI_{Effect} \cdot SI(t - SI_{delay})),$$

where  $SI_{Effect} \in [0, 1]$  is a parameter describing the efficacy of employed NPIs implying different adherence to the NPIs during the COVID-19 pandemic as well,  $SI_{delay}$ —time delay between government interventions enacting and their effect. See more details about Stringency Index in the Methods section.

The developed models may be accessed and simulated at <https://gitlab.sirius-web.org/covid-19/dde-epidemiology-model> via user interface or Jupyter Notebook.

**Simulation results.** The final version of the DDE based model was fitted to COVID-19 epidemic data in Germany and France from 01.01.2020 (model time  $t=0$ ) to 31.12.2021 (model time  $t=730$ ). All model parameter values are presented in Supplementary material 1 (Tables S1–S6). We have divided the overall time duration into four intervals or waves. Values of some model parameters were changed between waves to reflect changes in the pandemic progression. The most significant changes were made to the infection coefficient which causes a spike in new cases and reflects the spread of new more contagious SARS-Cov-2 variants<sup>33–36</sup>.

1. The First wave: This interval starts somewhere in January 2020. From this time point infected individuals started to arrive in the country in significant amounts. Patient zero in Germany entered the country on 20 January and was registered on 27 January<sup>37</sup>. We assumed in the model that import of infection to Germany began on 20 January. This import was ended on 16 March 2020 ( $t=76$ ), when the European Union as a whole announced the closure of all its external borders to non-citizens<sup>38</sup>. We assumed the import rate to be linearly increasing during that time period. Maximum number of infected individuals per day was estimated to be 500 individuals per day just before borders were closed.

For France the first case was identified on 24th January. However, individuals infected by SARS-CoV-2 were present as early as December 2019 according to some sources<sup>39,40</sup>. Unfortunately, we do not have data on how many infected individuals arrived in France or Germany before borders were closed on 16 March 2020. We kept the same number of persons per day for France as for Germany and estimated the start of infection import to France to be 15 January.

2. The second wave—starting from summer 2020 the number of new cases began to rise again implying the second epidemic wave in the region with many more registered cases. It may be attributed to a new European strain (EU1) emerging in both countries. However, its transmissibility is considered to be the same as for the original variant<sup>41</sup>. Another reason is relaxing anti epidemic restrictions which can be traced by lower levels of the Stringency Index. In both models it caused a second wave in accordance with existing statistics. Despite the fact that the number of cases is much higher than during the first wave, the number of hospitalized patients is roughly the same. In both models it was reflected by changing the fraction of severely ill among symptomatic patients  $H_F$ . New value was fitted according to the statistical data for patients in hospitals and ICU. New values were set in the model at time = 200 (20.07.2020) which agrees well with the appearance of the new strain in both countries according to the covariants.org web site. However, it should also be noted that the number of cases at the first pandemic wave could be dramatically underestimated due to limitations on testing compared to further deployment of the testing system. To estimate the correct number of initial cases data on seroprevalence during the first wave is required which is also limited.

3. The third wave—starting from early 2021 a new significant rise in numbers of cases begins in both countries. It may be connected with the spread of new strains of the virus. Indeed, new virus lineage with additional mutations in the spike region, B.1.1.7 strain<sup>42</sup>, was rapidly spreading in some European countries at this time period. This Alpha strain is much more contagious and has increased mortality rate according to the cohort study<sup>33–35</sup>. This was modeled by multiplying all probabilities to be infected upon contact by the same multiplier. The multiplier's value was implied to be the same for both countries and fitted to be 1.6. Thus, the probability of being infected upon contact is 60% larger for the new variant which is consistent with the estimated range of the transmissibility of the Alpha strain compared to the predecessor lineage<sup>43</sup>. Fraction of severely ill  $H_F$  was not changed at that time, as the previous value still agreed well with statistical data. Start of the third wave in the model was fitted for both countries on two different dates. Here it should be noted that the obtained profile of new cases for France does not agree well with the statistical data from ourworldindata.org.

4. The fourth wave starts in June 2021 and can be attributed to another B.1.617.2 virus variant (Delta) which is significantly more contagious than previous ones. New infection coefficient was fitted to be 2.3 times larger than for the Wuhan strain which agrees with estimates in the published data<sup>44,45</sup>. Fraction of severity ill  $H_F$  was fitted to be even less than for two previous strains to reflect a lowered ratio of hospitalized patients to registered. Despite the risk of hospital admission for COVID-19 was approximately doubled in patients with the Delta compared to the Alpha strain<sup>36</sup>, the overall hospital admissions involving COVID-19 in 2021 were significantly lower than in

2020 (e.g., see ONS data on COVID-19 latest insights: Hospitals. 9 June 2022). Apparently, the increased number of vaccinated people, protective effectiveness of the developed vaccines in preventing SARS-CoV-2 infections as well as a much greater proportion of the population cohort who is recovered from COVID-19 by the moment of Delta variant's emergence in 2021 ensured or can explain the ratio decline in this year compared to the first pandemic year. Once again, the time point at which new parameter values were introduced into the model was fitted for both countries for two different dates to obtain the required profiles of new cases.

Described four waves cover the first two years of the pandemic. As can be seen from simulation results (Figs. 3 and 4) the model accurately reproduces the reported new cases per week and total number of cases as well as the number of hospitalized patients on ICU and total deaths in each country over time of the pandemic for two years of the pandemic.

Sources used for deriving model parameters include: data on average number of contacts between individuals<sup>46</sup>, fatality rate on ICU<sup>47,48</sup>, transmissibility of the original virus strain<sup>49–51</sup>. Particular values and ranges of parameters may be found in Supplementary material 1.

**Automatic model generation for other countries.** Consequent development of our approach is automatic generation of epidemiological models for other countries (both European and non-European). Most of the model processes are fixed by fitting weighted sums of delays (see Methods). Because of that, the number of parameters whose values should be estimated is significantly lowered. That, in turn, allowed us to carry out estimation procedures automatically for other countries using statistical data provided by Our World in Data. It is worth to note that generated models are preliminary and further fine-tuning procedure is required to quantitatively reproduce observed epidemiological dynamics. However, it provides the fast generation of the initial version of epidemiological model for a certain country.

Automatic model generation goes as follows:

1. Copy of the base model is created.
2. Tabular data for a given country is derived from <https://covid.ourworldindata.org/data/owid-covid-data.csv>.
3. It is adapted to given country with following parameters:
  - a. Size of the population.
  - b. Number of vaccinated people each day.
  - c. Date when first case was registered  $T_0$ .
  - d. Start of infection import is set to  $T_0 - 15$  days.
  - e. Cancellation of the infection import (date when sufficient restrictions on borders were imposed) is set to  $T_0 + 30$  days.
  - f. Date when testing was started is set to 35 days (i.e. 4th February).
4. Parameters of the generated model are estimated so that weekly new cases simulated by the model agree with statistical data for a particular country. At the current stage, we estimated the model parameters at the first year of the epidemic in a certain country. Parameters, which values were estimated, are presented in Supplementary material 2 (Table S1).

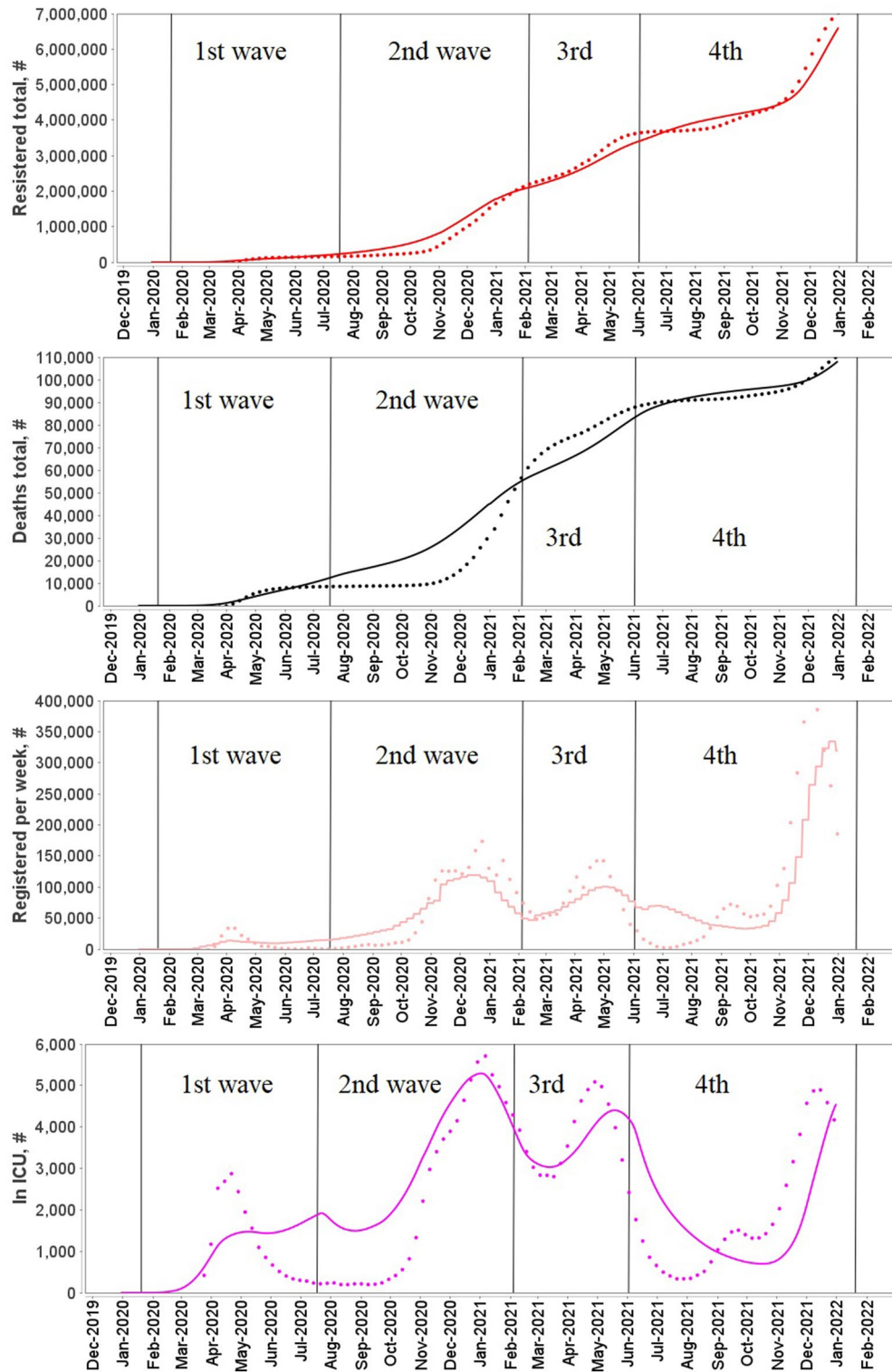
Simulation results of the automatically generated and fitted models for 12 countries are presented in Supplementary material 2 (Figures S1–S4). It is worth emphasizing that simulation results quite well agree with statistical data for most European countries, while optimization of the model describing COVID-19 epidemiology in Non-European countries like Brazil and Argentina was not able to adequately reproduce the observed trajectory of the epidemics. For example, one can see that the number of total cases in South Korea agrees well with real data, but the model failed to demonstrate the first two waves of pandemic. That is due to the fact that those waves are short and presented only by a few data points (2–3 weeks) and automatic parameter estimation ignores these small waves. Probably that manual model fitting and optimization are required in those cases. Of course, the generated models are not a final in silico tool to describe and predict epidemics even in European countries, but they may serve as a base for models of COVID-19 epidemic in corresponding countries.

## Discussion

We have proposed the methodology to overcome some shortcomings of the classic SEIR-based epidemiological models via the novel epidemiological model which utilizes DDEs to take into account different time scales of epidemiological processes and instantaneous splitting procedure to describe competing processes. The essential benefits of the developed model are:

1. Most of the epidemiological processes (symptoms onset, recovery, dying etc.) are described using kinetic laws with delayed arguments. These modeling processes can be fitted separately from the rest of the model and applied kinetic laws provide more precise reproduction of real properties of those processes than mass-action laws with a single parameter.
2. If a model has two or more competing transitions, a division into separate subpopulations removes their undesired mutual influence and allows to simulate fast and slow transitions with correct fractions of patients undergoing each transition.
3. Model parameters that are not parts of fitted processes described previously have direct mechanistic meaning (i.e. disease lethality, susceptibility, probability of different symptoms severity) and can be drawn from the statistics.

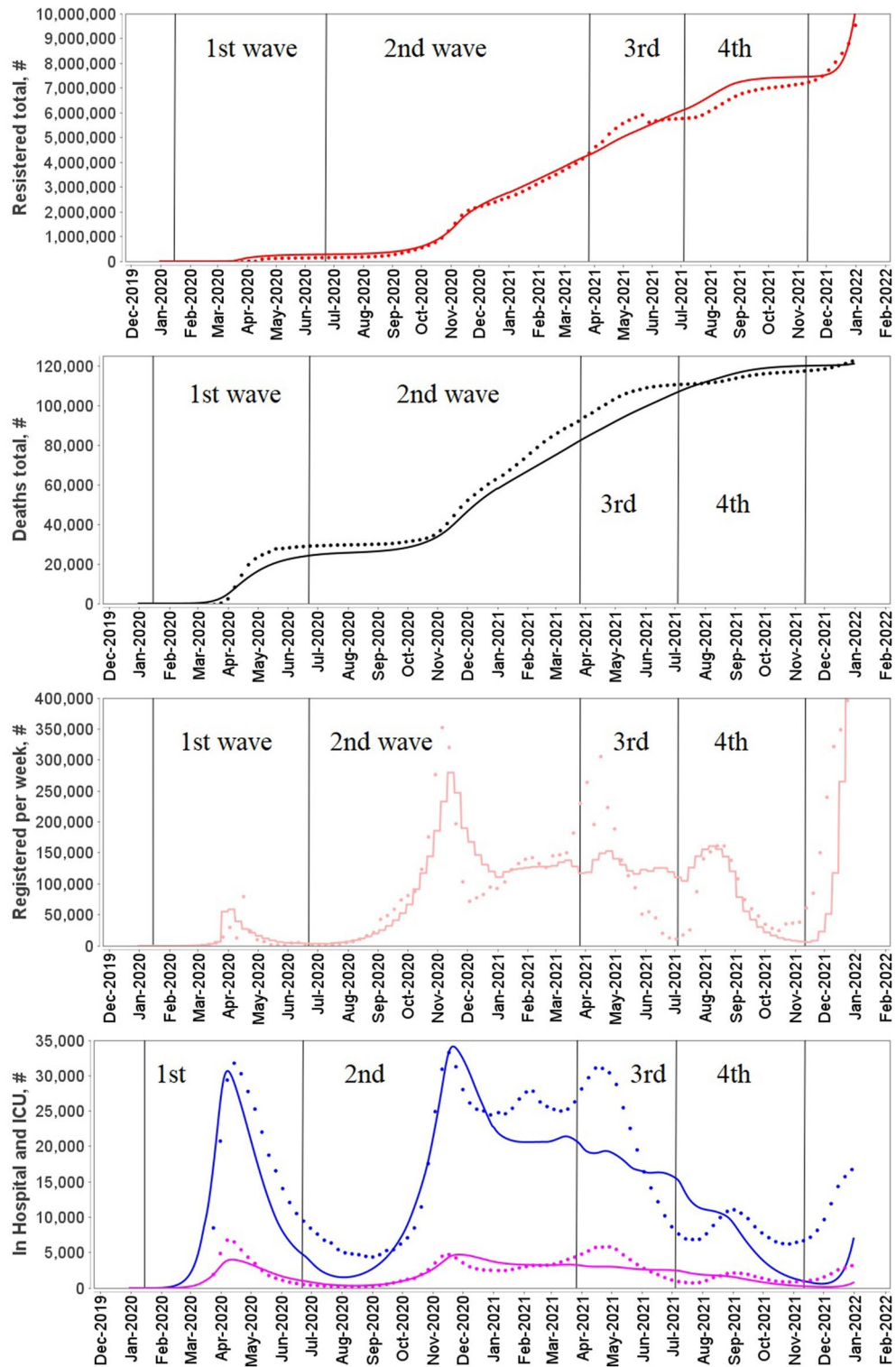




**Figure 3.** Simulation results and statistics from ourworldindata.org for Germany for 2020–2021 years.

- Overall decreasing of the parameters number makes them more identifiable as demonstrated in comparison between delay-based and classic models in this study. It also allows for more fast and simple adaptation of the model to other regions and countries.

Combination of these advantages makes the model more reliable with real properties of the pandemic and eliminates most of the abstract parameters usually used in SEIR-like models. It is worth noting that there are other studies addressing these issues using Erlang distribution<sup>52</sup> and delayed equations<sup>53</sup>. However, to our best



**Figure 4.** Simulation results and statistics from ourworldindata.org for France for 2020–2021 years.

knowledge, there are no publications demonstrating DDE-based approaches with weighted components of delayed arguments and instant transitions. Thus, we believe the proposed COVID-19 model and simulation results may be of interest to both experts in epidemiological modeling and a more general audience.

The final version of the DDE-based model describes many aspects of COVID-19 pandemic such as vaccination, loss of immunity over time, asymptomatic carriers, testing for virus, hospitalization for severely ill and intensive care required for critically ill. The model of course still has a number of abstract parameters reflecting government NPIs, conducted testing procedure, imposing quarantine for those who were detected as infected

with the virus and importing infected individuals into the modeling region or country. Particularly, parameters describing fractions of different symptoms' severity should be correctly attributed to the patient's age that requires extension of the model.

One of the crucial issues in the case of COVID-19 epidemiological modeling is to correctly transfer government NPIs (limit on mass gathering, lockdown, curfew, etc.) into the model parameters. One can easily see that introduction of a "social distance" multiplier to the infection rate and fitting its value to experimental data enables it to reproduce almost any observable epidemiological trajectory. In order to tackle this problem we tried to utilize the Stringency Index<sup>32</sup> instead of trying to fit the "social distance" factor over time of the pandemic in a certain country. However, we still have quite abstract aggregated numerical values of the indicator. Thus, further step for the model development in this direction is to use individual components or NPIs of the stringency index and attempt to assess their individual effect on the epidemic dynamics.

It should be noted that the main focus of the study is to present a novel, as far as we know, combination of delay-differential terms and instant processes which may be fitted separately from the rest of the SEIR model in order to simulate the pandemic. However, this approach and developed model for COVID-19 epidemic have some constraints and limitations. Firstly, the delay-differential version of the SEIR model like others assumes population homogeneity and it could be overcome only by application of agent-based modeling approach. Secondly, Covid-19 pandemic shows wave-like behavior in any country and the model analysis and fitting show that predictive power of this type of SEIR model like others is low and COVID-19 trajectory can not be predicted by the pre-fitted model without additional modifications taking into account an emergence of new viral strains with different epidemiological characteristics and consequent refitting of the model. So we associate each wave with appearance and spread of new more contagious variants of the virus and model it by changing infection parameters. However, it may be modeled more correctly using a modified version of the model with two or more similar modules where each of them is containing a full disease progressing scheme taking into account separate strains. This is also a part of our roadmap in the development of the model. The model fitting to statistical data for both countries demonstrated the decrease of the infection fatality rate during the COVID-19 epidemic which corresponds to early statements that the fatality rate of the Delta or B.1.617.2 variant of COVID-19, for example, is lower than the original variant. However, it might be caused by the age of unvaccinated people who were infected by the Delta virus strains and hospitalized with severe symptoms. According to the report published by Public Health England<sup>54</sup>, for instance, the majority of COVID-19 cases caused by the Delta variant were detected in people under 50 years old in the UK which are less likely to die from COVID-19 compared to those older than 50. So the current comparison of the case fatality rate of the B.1.617.2 variant with that of the wild-type virus is biased due to vaccination strategy in some European countries and age stratification of the population. This statistical misleading indicates the necessity to specify the developed model for each age group and integrate them in a more complex DDE model considering age distribution in a certain country. In addition to that, the vaccination scheme is oversimplified in the current version of the model and does not consider essential factors like two-stage and booster vaccinations, the vaccine efficiency against different strains (in regards to protective immunity to be infected and to show more serious symptoms), vaccination of asymptomatic cohort, different strategies or vaccine campaigns specific for particular country. So the roadmap for the extension and further model development includes the next biological aspects of the virus and epidemiology of the COVID-19 pandemic:

- Age-specific modules (infection and death rates, hospitalization and severity of the disease, B and T cell response) according to the statistical data<sup>55</sup>.
- Explicit strain emergence with other viral indicators like infectivity, resistance to neutralization, vaccine effectiveness in the population<sup>36,56–62</sup>.
- Waning B and T cell immunity and neutralization activity of specific antibodies<sup>63–66</sup>.
- Superspreading events<sup>67</sup>.
- Different vaccination strategies and its effectiveness against infection and severed outcomes for emerging viral strains.

## Methods

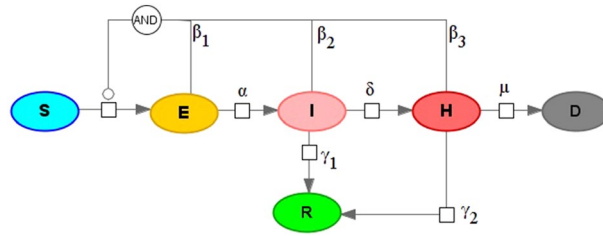
**SEIR-like model.** The overwhelming majority of SEIR-like models uses the mass action kinetic law (mainly, the first-order rate law) for transitions between different stages of a disease (e.g. between exposed and infectious periods). Explicit drawbacks of this approach are that:

1. Parameters of those reactions are quite abstract and can not be easily related to real biological characteristics of the virus.
2. The model fails to correctly represent processes delayed in time.

Here we will try to address those two problems. In the current study we use SBGN—Systems Biology Graphical Notation<sup>68</sup> for visual representation of mathematical models. Let's consider a SEIR-like model with two levels of symptom severity (Fig. 5).

Model equations:





**Figure 5.** SEIRHD model with mass-action kinetics.

$$\left\{ \begin{aligned} \frac{dS}{dt} &= -\frac{(\beta_1 \cdot E + \beta_2 \cdot I + \beta_3 \cdot H)}{(S + E + I + R + H + D)} \cdot S, \\ \frac{dE}{dt} &= \frac{(\beta_1 \cdot E + \beta_2 \cdot I + \beta_3 \cdot H)}{(S + E + I + R + H + D)} \cdot S - \alpha \cdot E, \\ \frac{dI}{dt} &= \alpha \cdot E - \gamma_1 \cdot I - \delta \cdot I, \\ \frac{dR}{dt} &= \gamma_1 \cdot I + \gamma_2 \cdot H, \\ \frac{dH}{dt} &= \delta \cdot I - \mu \cdot H - \gamma_2 \cdot H, \\ \frac{dD}{dt} &= \mu \cdot H. \end{aligned} \right. \tag{1}$$

With initial values:  $S(0) = S_0, E(0) = E_0, I(0) = R(0) = H(0) = D(0) = 0$ , where S—susceptible population, E—exposed (in incubation period) population, I—infected (with mild symptoms) population, H—infected (with severe symptoms), R—recovered population, D—dead individuals,  $\beta_1, \beta_2, \beta_3$ —infection rates for different contagious groups,  $\alpha$ —symptom onset rate,  $\delta$ —symptoms worsening rate,  $\mu$ —death rate,  $\gamma_1, \gamma_2$ —recovery rate (for mild and severe symptoms respectively).

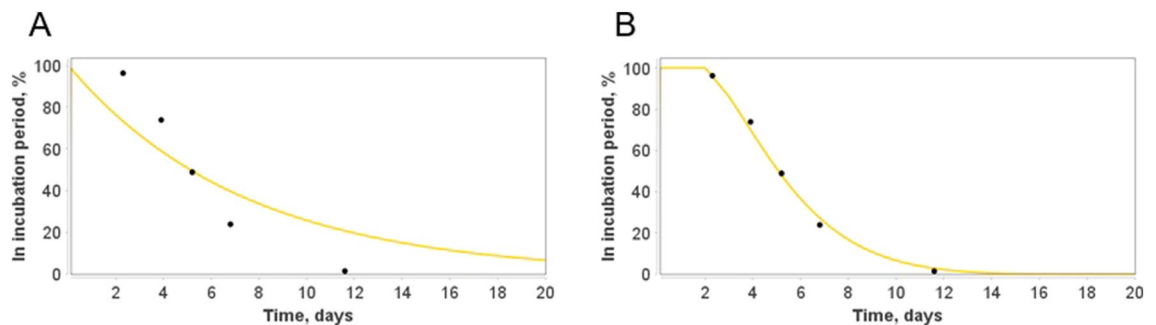
**Duration of processes.** A numerical value of the parameter  $\alpha$  in the model (1) is related to the median incubation period in the population. For example, if we set  $\alpha = \ln(2)/5.1$  then 50% of individuals who were exposed to the virus at time  $t=0$  will have become symptomatic at time  $t=5.1$  days.

$$\frac{dI}{dt} = -\frac{dE}{dt} = \frac{\ln(2)}{5.1}. \tag{2}$$

Distribution of incubation period in this model compared to the experimental data from<sup>28</sup> is presented in Fig. 6A. One can easily observe that it is inconsistent with statistical data on incubation period. For example, according to the first-order model, almost 10% of the infected have their symptom onset within one day of the infection which does not match the data.

Unfortunately, having only one parameter  $\alpha$  we can not fit the curve to this statistical data. This is also the case for other processes delayed in time with certain distribution of length.

A possible solution to overcome the issue is use of different forms of the kinetic law:



**Figure 6.** A comparison between two models of the incubation period: (A): based on the mass-action kinetics law (2). (B): using weighted sum of delays (4). Statistical data of the incubation period quantiles are taken from<sup>28</sup>.

$$\frac{dI}{dt} = -\frac{dE}{dt} = \sum_{i=0}^m k_i E(t - \Delta_i). \tag{3}$$

Herein, we use a weighted sum of the delayed number of exposed individuals. We have a  $2 \cdot m$  parameters which can be estimated to reproduce an experimental data. Typically,  $m = 1$  or  $m = 2$  is enough to comprehensively fit the data, keeping the number of parameters reasonably low.

For example, to fit the data from<sup>28</sup> we employ only two delays:

$$\frac{dI}{dt} = -\frac{dE}{dt} = 0.14E(t - 2) + 0.42E(t - 3). \tag{4}$$

Simulation results of the incubation period’s model demonstrating differences between theoretical curves obtained using two methodologies are presented in Fig. 6.

Another benefit of the time-delay based approach is an opportunity to reproduce an experimental data on diverse epidemiological processes (incubation period, recovery, worsening of symptoms from mild to severe and others) once and separately from the rest of the model structure based on the known distributions of duration of those processes for particular infectious disease only.

**Competing processes.** Another opportunity to improve the original model and bring it closer to reality is to consider different possible transitions from the same subgroup. For example, infected patients may either recover or progress to severe symptoms (Fig. 7). In that case parameters  $\delta$  and  $\gamma_1$  are related not only to durations of corresponding processes but to recovery rate for mild symptomatic (or fraction of severe symptomatic among mild symptomatic). An issue with two alternatives arises when the fast process has less probability and therefore smaller fraction of patients involved in this direction of the infectious process. According to the statistics<sup>29</sup>, the process of worsening of symptoms (median time is 5 days) is faster than recovery (median time equals to 14 days) implying that  $\delta$  value should be larger than  $\gamma_1$  value. As in the previous subsection, we may set these parameters as  $\delta = \ln(2)/5, \gamma_1 = \ln(2)/14$ . However, in that case (Fig. 7) the model will show that the fraction of patients who transit to the severe symptoms is larger than the fraction recovered which is not the case in reality<sup>69</sup>.

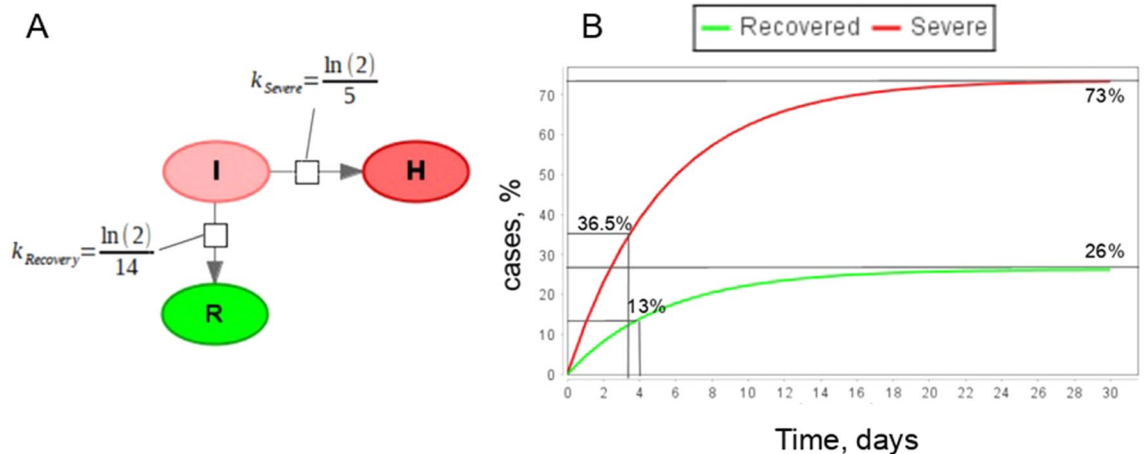
The possible solution to overcome the discrepancy is to consider those two processes separately. Firstly, patients will instantaneously transit from “Symptomatic” to “Symptomatic who will recover in future” ( $I_R$ ) and then from  $I$  to  $R$ . In similar way another fraction of symptomatic patient will instantaneously transit to “Symptomatic who will need hospitalization” ( $I_H$ ) and only afterwards from  $I$  to  $H$ . The updated model is presented in Fig. 8 and corresponding model equations are:

$$\begin{cases} \frac{dI_O}{dt} = -k_{Large} \cdot I_O, \\ \frac{dI_R}{dt} = k_{Large} \cdot F_R \cdot I_O, \\ \frac{dI_H}{dt} = k_{Large} \cdot F_H \cdot I_O. \end{cases} \tag{5}$$

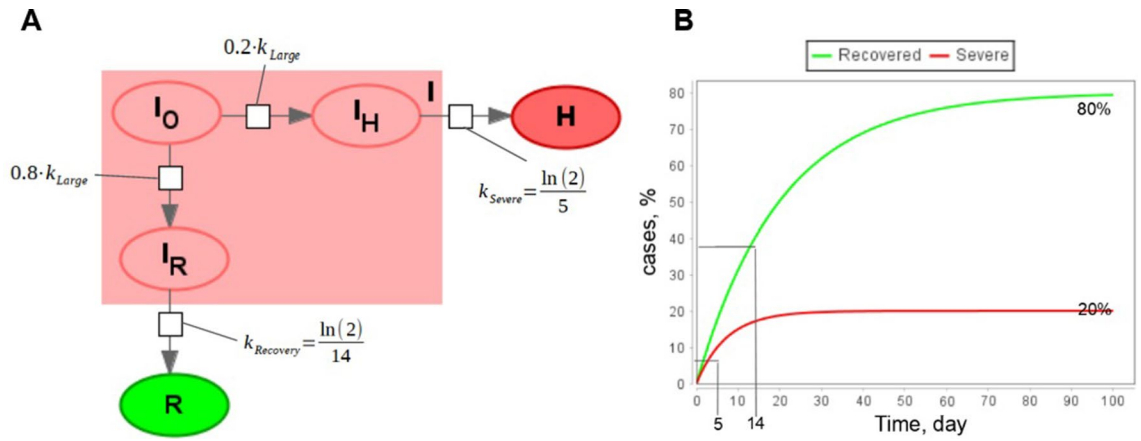
where  $F_R = 0.2$ —a fraction of infectious individuals who will not have worse symptoms.  $F_H = 0.8$ —a fraction of infectious individuals who will have worse symptoms.  $F_R + F_H = 1$ .  $k_{Large}$ —constant value which is large enough to render reactions instant.

The total number of individuals with mild symptoms is calculated as a sum of all subgroups:

$$I = I_O + I_R + I_H$$



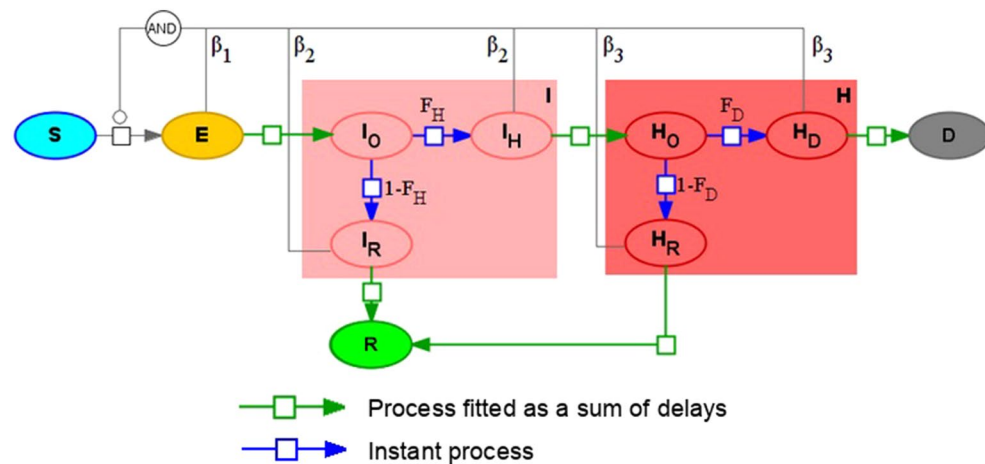
**Figure 7.** The simple mass-action model with two competing processes. (A): SBGN representation of the model, here I—infected individuals, R—recovered, H—patients with severe symptoms; (B): Simulation results of the model.



**Figure 8.** A part of the epidemiological model with alternative competing processes. (A): Visual representation in SBN format; (B): Simulation results of the model for fractions of recovered and patients with severe symptoms.

This technique can be combined with delayed equations described in the previous section. Thus, we can construct a version of the model (1) taking into account the duration of some infectious processes and the existence of competing processes. The final version of the model is presented in Fig. 9.

Advantage of the updated model is that instead of 8 parameters which do not explicitly correspond to real characteristics of infectious processes and has to be fitted to experimental data we have only three parameters, two fraction parameters:  $F_H$ —fraction of symptomatic individuals with severe symptoms,  $F_D$ —disease lethality that can be drawn from statistical data and processes that are fitted to experimental data separately from the rest of the model. Comparison is given in Table 1.



**Figure 9.** SEIR-like model with instant and delay-based processes (2).

Modeled process	Parameters	
	SEIRHD model (1)	Delay-based model (2)
Exposure to the virus	Infection rates $\beta_1, \beta_2, \beta_3$	Infection rates $\beta_1, \beta_2, \beta_3$
Incubation period	$\alpha$	Fitted process
Recovery with mild symptoms	$\gamma_1$	Fraction of severe symptomatic $F_H$ , fitted process
Progress to severe symptoms	$\delta$	Fraction of severe symptomatic $F_H$ , fitted process
Recovery with severe symptoms	$\gamma_2$	Disease lethality $F_D$ , fitted process
Death	$\mu$	Disease lethality $F_D$ , fitted process

**Table 1.** Comparison of parameters in classic SEIRHD model and delay-based model.

**Comparison of the model parameters.** Advantage of the updated model is that only three epidemiologically interpretable parameters—infection rates  $\beta_1, \beta_2, \beta_3$  have to be fitted to experimental data instead of 8 parameters  $\alpha, \beta_1, \beta_2, \beta_3, \gamma_1, \gamma_2, \delta, \mu$  which do not explicitly correspond to real characteristics of infectious processes. We also have two fraction parameters which values may be explicitly derived from statistics:  $F_H$ —fraction of symptomatic individuals with severe symptoms,  $F_D$ —disease lethality. All other parameters of the new model are fitted preliminary to experimental data on corresponding processes separately from the rest of the model. Comparison is given in Table 1. Decreasing of parameters number also leads to better identifiability of remaining parameters. In order to demonstrate that we fitted two type of the SEIR model (classic and delay-differential) to data on the first 180 days of the pandemic in Germany in 2020. We simulate a start of the pandemic by the simplistic way and use two additional parameters:  $Start$ —day, when the infection was imported to the country and  $E_{start}$ —a number of individuals in incubation period imported to the country at day  $Start$ . Results of the parameter estimation and parameter identifiability analysis for both delay-based and classic SEIR models are presented in Supplementary material 3, more details may be also found in corresponding Jupyter Notebook at <https://gitlab.sirius-web.org/covid-19/dde-epidemiology-model>.

According to the analysis for delay-based SEIR model:  $\beta_1$  is identifiable parameter (i.e. if its value is changed from estimated  $\beta_1 = 0.165$  and fixed, then the model can not be successfully fitted by changing other parameters values);  $\beta_2$ —is non-identifiable in very small range [0, 0.0026] (i.e. its value may be anywhere in this range), value of  $\beta_3$ —is also non-identifiable in quite a small range [0, 0.029]. We can see that all three parameters can be identified quite precisely.

However, the parameter identifiability analysis has been demonstrated that for classic SEIR model:  $\beta_1$  is not-identifiable in the range [0, 9.92],  $\beta_2, \beta_3$  are not identifiable in the range [0, 10]. If we fix the value of one of these parameters anywhere in that range we may fit the model by changing values of other parameters.  $Start$  is identifiable, where the identified value is day 83 (23.03.2020), while  $E_{start}$  is non-identifiable in the range [10570, 2000] and  $\alpha$  is non-identifiable in the range [0.57, 10]. Other parameters are also not-identifiable in large ranges.

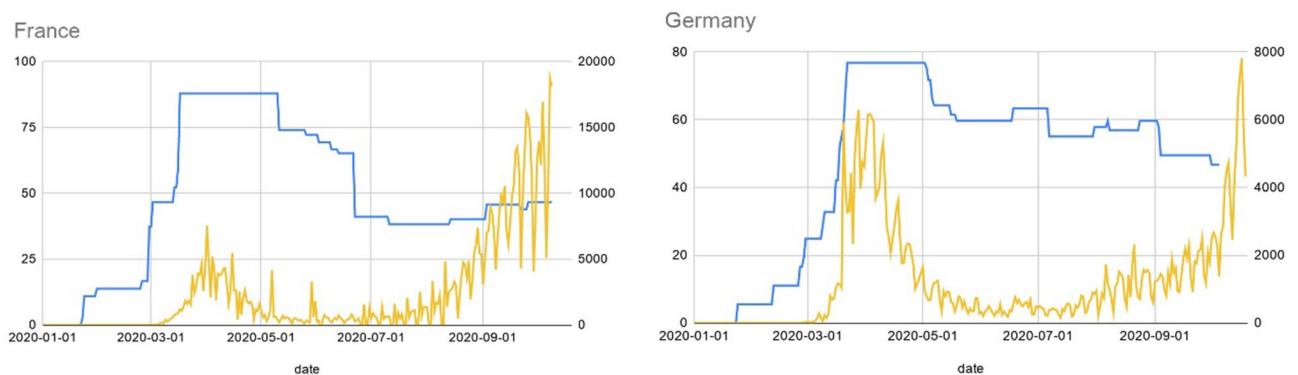
Thus, the delay-differential modelling approach not only provides a smaller number of epidemiologically interpretable parameters, but also improves the identifiability of the model parameters compared to classic SEIR model.

**Initial model.** As a basis for our model we used the previously created SEIR model<sup>16</sup> of the COVID-19 epidemic. This model differs from the most SEIR models by differentiating between tested and non-tested infected subjects. It was created in the Systems Biology software COPASI<sup>70</sup> which allows one to specify the kinetics of the process mechanistically. COPASI translates these specifications into differential equations which it integrates either as a function of time, or by requiring steady state. The software honors restrictions as specified in terms of algebraic equations and ‘events’ which instantaneously change numeric values of the model parameters triggered by logical expressions transiting from “false” to “true”. COPASI models are SBML compatible, and can be exported into the format. The latter greatly facilitates model reuse and reproduction.

**Data sources.** Statistical data for Germany and France was taken from Our World In Data web site (<https://ourworldindata.org/>). This web-portal provides the data on the total number of cases, new cases each day, number of hospitalized, transferred to ICU patients, vaccinated individuals and total number of deaths.

To take into account statistical data on government actions in the model we employed the Stringency Index developed by the Blavatnik School of Government of the University of Oxford<sup>32</sup> which incorporates data on 10 different measures corresponded to: C1—School closing, C2—Workplace closing, C3—Public events cancellation, C4—Restriction on gatherings, C5—Public transport closing, C6—Stay at home requirements, C7—Internal movement, C8—International movement, H1—public information campaigns. Comparison between Stringency index and new registered cases is provided in Fig. 10.

**BioUML platform.** BioUML (<http://www.biouml.org>) used in the study is an integrated Java platform for modeling of biological systems<sup>71</sup>. It supports different mathematical formulations for the model development including ODE, delay-based, algebraic systems, discrete events, agent-based, and stochastic modeling. The



**Figure 10.** Stringency of government measures (blue) and new reported cases per day (yellow).

platform also incorporates a module for the automatic and manual parameter fitting to an experimental data. Models developed in the BioUML are based on main standards in systems biology: (1) SBML—Systems Biology Markup Language<sup>72</sup> for mathematical description and (2) SBGN for visual representation. A model can be built and edited in the platform as a visual diagram (e.g. in SBGN notation) based on which a Java code is generated for model simulations. Additionally, BioUML is integrated with Jupyter hub (<https://jupyter.org/>) for interactive data and model analysis as well as an essential and user-friendly tool for reproducibility of the simulation results.

**Parameter identifiability.** Assessing of the parameter identifiability was conducted using a method implemented in the platform<sup>73</sup>. Identifiability analysis examines whether a set of model parameter values can be uniquely estimated from a given model and data set. According to the methodology, the agreement of experimental data with the observables predicted by the model is measured by an objective function, commonly the weighted sum of squared errors. The analysis goes as follows: one of the parameters is selected, its value is fixed and the model is fitted to observed data using other parameters. Then the fixed parameter value is changed in order to find a range of values for a given parameter in which the model could not be fitted to observed data. An area in which the model may be fitted regarding a given parameter value (inside this region) is the range in which this parameter is unidentifiable (its value cannot be uniquely found considering current data, model and set of parameters for estimation). Then the procedure repeats for the next model parameter.

### Data availability

The data that support the findings of this study are available from <https://ourworldindata.org/>. Datasets selected for particular regions are also available through the web interface of BioUML software at <https://gitlab.sirius-web.org/covid-19/dde-epidemiology-model>. The developed models are available through the web interface of BioUML software at <https://gitlab.sirius-web.org/covid-19/dde-epidemiology-model>. Models are available both through visual representation in the platform and in Jupyter notebooks which allow users to reproduce simulation results and figures presented in this study.

Received: 23 June 2022; Accepted: 3 August 2023

Published online: 18 August 2023

### References

1. Metcalf, C. J. E., Morris, D. H. & Park, S. W. Mathematical models to guide pandemic response. *Science* **369**(6502), 368–369 (2020).
2. Adam, D. Special report: The simulations driving the world's response to COVID-19. *Nature* **580**(7802), 316–319 (2020).
3. Ciuriak, D. & Fay, R. The critical numbers game: How models can inform the pandemic policy response from lockdown to reboot. Opinion, Centre for International Governance Innovation (2020).
4. Cobey, S. Modeling infectious disease dynamics. *Science* **368**(6492), 713–714 (2020).
5. Kucharski, A. J. *et al.* Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect. Dis.* **20**(5), 553–558 (2020).
6. Kermack, W. O. & McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. A Contain. Math. Phys. Character* **115**(772), 700–721 (1927).
7. Atkeson, A. What will be the economic impact of COVID-19 in the US? Rough estimates of disease scenarios (No. w26867). *NBER Work Pap Ser* (2020).
8. Calafiore, G. C., Novara, C. & Possieri, C. A time-varying SIRD model for the COVID-19 contagion in Italy. *Annu. Rev. Control* **50**, 361–372 (2020).
9. Fanelli, D. & Piazza, F. Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos Solitons Fractals* **134**, 109761 (2020).
10. Zhang, Y. *et al.* Epidemiological investigation on a cluster epidemic of COVID-19 in a collective workplace in Tianjin. *Zhonghua liu xing bing xue za zhi = Zhonghua liuxingbingxue zazhi* **41**(5), 649–653 (2020).
11. Banerjee, M., Tokarev, A. & Volpert, V. Immuno-epidemiological model of two-stage epidemic growth. *Math. Model. Nat. Phenom.* **15**, 27 (2020).
12. Li, M. Y. & Muldowney, J. S. Global stability for the SEIR model in epidemiology. *Math. Biosci.* **125**(2), 155–164 (1995).
13. Prem, K. *et al.* The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: A modelling study. *Lancet Public Health* **5**(5), e261–e270 (2020).
14. Yang, Z. *et al.* Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *J. Thorac. Dis.* **12**(3), 165 (2020).
15. Ferguson, N. M. *et al.* Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College COVID-19 Response Team. Imperial College COVID-19 Response Team 20 (2020).
16. Westerhoff, H. V. & Kolodkin, A. N. Advice from a systems-biology model of the corona epidemics. *npj Syst. Biol. Appl.* **6**(1), 1–5 (2020).
17. Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H. & Lipsitch, M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**(6493), 860–868 (2020).
18. Chinazzi, M. *et al.* The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**(6489), 395–400 (2020).
19. Tuomisto, J. T. *et al.* An agent-based epidemic model REINA for COVID-19 to identify destructive policies. *medRxiv* <https://doi.org/10.1101/2020.04.09.20047498> (2020).
20. Cooke, K. L. & Van Den Driessche, P. Analysis of an SEIRS epidemic model with two delays. *J. Math. Biol.* **35**(2), 240–260 (1996).
21. Martcheva, M. *An Introduction to Mathematical Epidemiology* Vol. 61 (Springer, New York, 2015).
22. Shayak, B., Sharma, M. M., Rand, R. H., Singh, A. & Misra, A. N. O. P. A Delay differential equation model for the spread of COVID-19. *Int. J. Eng. Res.* **10**(103), 1–13 (2020).
23. Götz, T. & Heidrich, P. Early stage COVID-19 disease dynamics in Germany: models and parameter identification. *J. Math. Ind.* **10**(1), 1–13 (2020).
24. Menendez, J. Elementary time-delay dynamics of COVID-19 disease. *medRxiv* <https://doi.org/10.1101/2020.03.27.20045328> (2020).
25. Sharma, N., Verma, A. K. & Gupta, A. K. Spatial network based model forecasting transmission and control of COVID-19. *Phys. A: Stat. Mech. Appl.* **581**, 126223 (2021).
26. Utamura, M., Koizumi, M. & Kirikami, S. Isolation considered epidemiological model for the prediction of COVID-19 trend in Tokyo, Japan: Numerical Study. *JMIR Public Health Surveill* (2020).



27. Oran, D. P. & Topol, E. J. The proportion of SARS-CoV-2 infections that are asymptomatic : A systematic review. *Ann. Intern. Med.* **174**(5), 655–662 (2021).
28. Lauer, S. A. *et al.* The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application. *Ann. Intern. Med.* **172**(9), 577–582 (2020).
29. Boëlle, P. Y. *et al.* Trajectories of hospitalization in COVID-19 patients: an observational study in France. *J. Clin. Med.* **9**(10), 3148 (2020).
30. Baraniuk, C. How long does covid-19 immunity last?. *BMJ* **373**, n1605 (2021).
31. Radbruch, A. & Chang, H. D. A long-term perspective on immunity to COVID. *Nature* **595**, 359–360 (2021).
32. Hale, T. *et al.* A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nat. Hum. Behav.* **5**(4), 529–538 (2021).
33. Tao, K. *et al.* The biological and clinical significance of emerging SARS-CoV-2 variants. *Nat. Rev. Genet.* **22**, 757–773 (2021).
34. Jones, T. C. *et al.* Estimating infectiousness throughout SARS-CoV-2 infection course. *Science* **373**(6551), eabi5273 (2021).
35. Challen, R. *et al.* Risk of mortality in patients infected with SARS-CoV-2 variant of concern 202012/1: Matched cohort study. *BMJ* **372**, n579 (2021).
36. Sheikh, A., McMenamin, J., Taylor, B. & Robertson, C. SARS-CoV-2 Delta VOC in Scotland: demographics, risk of hospital admission, and vaccine effectiveness. *Lancet* **397**, 2461–2462 (2021).
37. Böhmer, M. M. *et al.* Investigation of a COVID-19 outbreak in Germany resulting from a single travel-associated primary case: A case series. *Lancet Infect. Dis.* **20**(8), 920–928 (2020).
38. European Commission. “COVID-19: Temporary Restriction on Non-Essential Travel to the EU.” *European Commission*. accessed 16 March 2020. <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020DC0115&from=EN>. (2020)
39. Deslandes, A. *et al.* SARS-CoV-2 was already spreading in France in late December 2019. *Int. J. Antimicrob.* **55**(6), 106006 (2020).
40. Carrat, F. *et al.* Evidence of early circulation of SARS-CoV-2 in France: findings from the population-based “CONSTANCES” cohort. *Eur. J. Epidemiol.* **36**(2), 219–222 (2021).
41. Hodcroft, E. B. *et al.* Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature* **595**(7869), 707–712 (2021).
42. Frampton, D. *et al.* Genomic characteristics and clinical effect of the emergent SARS-CoV-2 B. 1.1.7 lineage in London, UK: A whole-genome sequencing and hospital-based cohort study. *Lancet Infect. Dis.* **21**(9), 1246–1256 (2021).
43. Davies, N. G. *et al.* Estimated transmissibility and impact of SARS-CoV-2 lineage B. 1.1.7 in England. *Science* **372**(6538), eabg3055 (2021).
44. Liu, Y. & Rocklöv, J. The reproductive number of the Delta variant of SARS-CoV-2 is far higher compared to the ancestral SARS-CoV-2 virus. *J. Travel Med.* **28**(7), taab124 (2021).
45. Despres, H. W. *et al.* Measuring infectious SARS-CoV-2 in clinical samples reveals a higher viral titer: RNA ratio for Delta and Epsilon versus Alpha variants. *PNAS* **119**(5), e2116518119 (2022).
46. Del Fava, E. *et al.* The differential impact of physical distancing strategies on social contacts relevant for the spread of COVID-19. *BMJ Open* **11**(10), e050651 (2021).
47. Karagiannidis, C., Windisch, W., McAuley, D. F., Welte, T. & Busse, R. Major differences in ICU admissions during the first and second COVID-19 wave in Germany. *Lancet Respir. Med.* **9**(5), e47–e48 (2021).
48. Rimmelé, T., Pascal, L., Polazzi, S. & Duclos, A. Organizational aspects of care associated with mortality in critically ill COVID-19 patients. *Intensive Care Med.* **47**(1), 119–121 (2021).
49. Li, R. *et al.* Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**(6490), 489–493 (2020).
50. Park, S. W. *et al.* Reconciling early-outbreak estimates of the basic reproductive number and its uncertainty: Framework and applications to the novel coronavirus (SARS-CoV-2) outbreak. *J. R. Soc. Interface* **17**(168), 20200144 (2020).
51. Xu, H. *et al.* Basic reproduction number of the 2019 novel coronavirus disease in the major endemic areas of China: a latent profile analysis. *Public Health Front.* **9**, 575315 (2021).
52. Arino, J. & Portet, S. A simple model for COVID-19. *Infect. Dis. Model.* **5**, 309–315 (2020).
53. Devipriya, R., Dharmodharavadhani, S., & Selvi, S. SEIR model for COVID-19 epidemic using delay differential equation. In *Journal of Physics: Conference Series*, vol. 1767, 012005 (2020).
54. Public Health England, SARS-CoV-2 variants of concern and variants under investigation in England. *Technical briefing 18*. URL [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/1001358/Variants\\_of\\_Concern\\_VOC\\_Technical\\_Briefing\\_18.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1001358/Variants_of_Concern_VOC_Technical_Briefing_18.pdf). [Online] (Accessed: 22 July 2021) (2021).
55. Yang, F. *et al.* Shared B cell memory to coronaviruses and other pathogens varies in human age groups and tissues. *Science* **372**(6543), 738–741 (2021).
56. Chen, R. E. *et al.* Resistance of SARS-CoV-2 variants to neutralization by monoclonal and serum-derived polyclonal antibodies. *Nat. Med.* **27**(4), 717–726 (2021).
57. Davies, N. G. *et al.* Increased mortality in community-tested cases of SARS-CoV-2 lineage B. 1.1. 7. *Nature* **593**(7858), 270–274 (2021).
58. Faria, N. R. *et al.* Genomics and epidemiology of the P. 1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* **372**(6544), 815–821 (2021).
59. Jalkanen, P. *et al.* COVID-19 mRNA vaccine induced antibody responses and neutralizing antibodies against three SARS-CoV-2 variants. *Nat. Commun.* **12**, 3991 (2021).
60. Volz, E. *et al.* Assessing transmissibility of SARS-CoV-2 lineage B. 1.1. 7 in England. *Nature* **593**(7858), 266–269 (2021).
61. Washington, N. L. *et al.* Emergence and rapid transmission of SARS-CoV-2 B. 1.1. 7 in the United States. *Cell* **184**(10), 2587–2594 (2021).
62. Zhou, D. *et al.* Evidence of escape of SARS-CoV-2 variant B. 1.351 from natural and vaccine-induced sera. *Cell* **184**(9), 2348–2361 (2021).
63. Dan, J. M. *et al.* Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* **371**(6529), eabf4063 (2021).
64. Harrington, W. E. *et al.* Rapid decline of neutralizing antibodies is associated with decay of IgM in adults recovered from mild COVID-19. *Cell Rep.* **2**(4), 100253 (2021).
65. Sherina, N. *et al.* Persistence of SARS-CoV-2-specific B and T cell responses in convalescent COVID-19 patients 6–8 months after the infection. *Med* **2**(3), 281–295 (2021).
66. Zuo, J. *et al.* Robust SARS-CoV-2-specific T cell immunity is maintained at 6 months following primary infection. *Nat. Immunol.* **22**(5), 620–626 (2021).
67. Lemieux, J. E. *et al.* Phylogenetic analysis of SARS-CoV-2 in Boston highlights the impact of superspreading events. *Science* **371**(6529), eabe3261 (2021).
68. Novere, L. *et al.* The systems biology graphical notation. *Nat. Biotechnol.* **27**(8), 735–741 (2009).
69. WHO Situation Report 6 March. [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200306-sitrep-46-covid-19.pdf?sfvrsn=96b04adf\\_4](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200306-sitrep-46-covid-19.pdf?sfvrsn=96b04adf_4) (2020).
70. Hoops, S. *et al.* COPASI—A complex pathway simulator. *Bioinformatics* **22**(24), 3067–3074 (2006).
71. Kolpakov, F. *et al.* BioUML—Towards a universal research platform. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkac286> (2022).

72. Hucka, M. *et al.* The systems biology markup language (SBML): language specification for level 3 version 2 core. *J. Integr. Bioinform.* **15**(1), 20170081 (2018).
73. Raue, A. *et al.* Identifiability and observability analysis for experimental design in nonlinear dynamical models. *Chaos Interdiscip. J. Nonlinear Sci.* **20**(4), 045105 (2010).

### Acknowledgements

The study was supported by the RFBR (#20-04-60355 and by the Ministry of Science and Higher Education of the Russian Federation (Agreement 075-10-2021-093, Project CMB-RND-2123).

### Author contributions

K.I.N. designed and implemented the model with help and consultation from A.I.R. and K.F.A. K.I.N. and A.I.R. drafted the manuscript. K.F.A. coordinated the project. All authors read and approved the final manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-40008-9>.

**Correspondence** and requests for materials should be addressed to I.N.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023