Simulation of non-pharmaceutical interventions in an agent based epidemic model

Petra Vidnerová[1,9], Roman Neruda[1,9], Gabriela Suchopárová[1,9], Luděk Berec[3,9], Tomáš Diviák[4,9], Aleš Kuběna[2,9], René Levínský[5,9], Josef Šlerka[6,9], Martin Šmíd[2,9], Jan Trnka[7,9], Vít Tuček[8,9], Karel Vrbenský[2,9], and Milan Zajíček[2,9]

[1] The Czech Academy of Sciences, Institute of Computer Science
petra@cs.cas.cz
[2] The Czech Academy of Sciences, Institute of Information Theory and Automation
[3] Institute of Mathematics, Faculty of Science, University of South Bohemia and The Czech Academy of Sciences, Biology Centre, Institute of Entomology
[4] Department of Criminology and Mitchell Centre for Social Network Analysis, School of Social Sciences, University of Manchester
[5] CERGE-EI
[6] New Media Studies, Faculty of Arts, Charles University
[7] Department of Biochemistry, Cell and Molecular Biology, Third Faculty of Medicine, Charles University
[8] Department of Mathematics, University of Zagreb
[9] Centre for Modelling of Biological and Social Processes
WWW home page: http://bisop.eu

Abstract: The standard SEIR equation-based models represent the state-of-the-art approach in epidemiological modelling. Their drawbacks include unrealistic infection-related contact estimates and difficulties in modelling non-pharmaceutical interventions, such as contact reductions or partial closures.

In this paper, we present our agent-based model that addresses the above-mentioned issues. It works with a population of individuals (agents) and their contacts are modelled as a multi-graph social network according to real data based on a Czech county. Custom algorithmic procedures simulating testing, quarantine and partial closures of various contact types are implemented.

The model can serve as a tool for relative comparison of the efficacy of various policies. It was also used for a study comparing various interventions in Czech primary and secondary schools, using a graph based on real data from a selected Czech school.

1 Introduction

Mathematical models play an important role in a variety of scientific fields, including epidemiology. Mathematical modelling has been an integral part of epidemiology for more than 100 years. Epidemiological models serve many different purposes, namely as a tool for hypothesis verification, explanation of observed data, understanding basic principles of infectious dynamics, prediction of the future development and understanding the current one, calculation of fundamental epidemiological metrics.

During the current world-wide pandemic the interest in epidemiological modelling significantly increased not only in the scientific community, but also in the general public. The ability to model and predict the development of the epidemic became important from day to day.

There exists a variety of epidemic models, including well established S(E)IR models[14, 5]. Our focus is on modelling the impact of non-pharmaceutical interventions. The non-pharmaceutical interventions and epidemiological measures influence the development of the epidemic significantly. Therefore, a proper epidemic model should reflect these interventions and must be able to model them and their impact on epidemic. Such a model may then become an important tool in choosing most efficient policies.

In this paper we present our agent based model that was designed for the purpose of comparing various interventions, measures and policies. We focus on the simulation of non-pharmaceutical interventions, such as partial closures, contact restrictions and quarantines. More details on the model itself can be found in our preprint [6].

The interventions we simulate cover protective measures and contact restrictions. By protective measures we understand masks, stronger hygiene, and general cautiousness. Protective measures decrease the probability of transmission of the disease when the infectious contact happens. On the other hand, the contact restrictions decrease the probability that such contacts are realised. Contact restrictions can be either flat, i. e. whole public places are closed (school closures, shop closures, etc.), or individual. Individual contact restrictions cover isolation or quarantine of single individuals. The effective contact restriction on individual level requires to actively search for individuals that were in contact with infectious ones. This process is known as *contact tracing*.

All mentioned interventions play an important role in the epidemic and the model should be able to reflect them.

Some of the available epidemiological models posses mechanisms for modelling some of the mentioned interventions including contact tracing [17, 13, 12, 11, 8].

The paper is organised as follows. In the next section we briefly explain the framework of our model and its main properties. In Section 3 the principles of simulation of individual interventions are described. Section 4 brings discussion of the usage of the model, its capabilities and

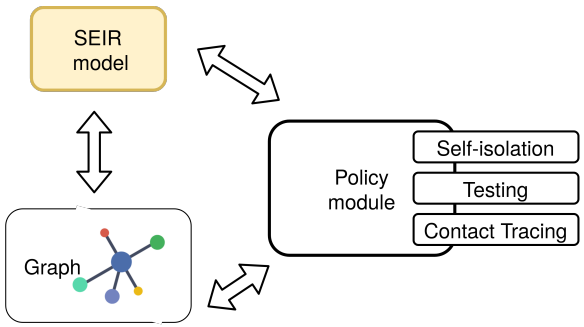

Figure 1: The structure of our model.

limitations. The final section is a conclusion.

2 Agent Based Epidemic Model

Our model can be classified as a multi agent model. A key property of agent models is that they are built in a bottom-up manner. We learn about the behaviour of the system as a whole from the detailed description of its components by means of agents and their interactions by computer simulation. This is completely different approach than in classical compartment models, where the whole system is described by set of equations.

The core of our model framework is formed by three modules. They are depicted in the Fig. 1. Namely, they are the SEIR model, the contact graph and the policy module. Let us describe them briefly.

SEIR model

The SEIR model module is the base part of our model. It is a standard epidemiological model of the S(E)IR type [5]. It works with a population of individuals (agents), each individual being in one of possible states. The basic states are S (*susceptible*, healthy individuals), E (*exposed*, infected but not yet infectious), I (*infectious*), R (*recovered*). Our model, in addition, distinguishes the asymptomatic, presymptomatic and symptomatic individuals; and an infectious and post-infectious phase. Therefore it works with 10 states in total. Additionally, each individual keeps a flag whether it was detected. An example of possible node life cycle is $S \to E \to I_a \to I_s \to J_s \to R$, where I_a stands for infectious presymptomatic, I_s for infectious symptomatic, J for not infectious, but yet positive.

One iteration of the model corresponds to one day, i. e. an individual can change its state once a day. The transitions between the states, in other words times for which the individual stays in each state, are given by the parameters of the disease (in this case COVID-19). For details on model parameters used for COVID-19 simulations see [6].

The exception is the transition $S \to E$ that depends both on the infectiousness of the disease modelled (a global model parameter β; note that β has a different meaning than in typical SEIR models) and the contact graph. This transition is also the one influenced by non-pharmaceutical interventions.

Contact Graph

While the majority of epidemiological models use synthetic population (if any), our model uses a multi-graph that is based on real data. The graph is built as a model of a Czech county (a town together with its surrounding villages). Great attention was given to the graph construction so as it is as realistic as possible. Main data sources used include Czech Statistical Office [10] (population data collected in 2011), State Administration of Land Surveying and Cadastre [3], Open Street Map [2] and publicly available database of schools [18]. Contacts between individuals are established based on sociological studies by PAQ [19] and MEDIAN [16, 15]. The contacts between friends were generated by a modified version of Barabasi-Albert algorithm [4]. The probabilities of contacts were automatically optimised to fit Prem matrices [20] (see Fig. 3).

The resulting graph has about 56 thousands nodes (representing agents) and 2.8 millions edges (contacts between agents). Nodes have a list of attributes, including sex, age and economic activity. An example of an individual (and a node in a graph) can be a 43 years old teacher, who lives with his wife and three children in a family house. He has 12 friends, who he regularly meets, most of them are of the similar age. Every day he commutes to work and regularly visits his parents in a close village. All these activities have to be reflected by contact edges in a graph.

The detailed explanation of the graph construction is out of the scope of this paper and can be found in [6].

Formally, the *contact graph* is a multi-graph $G = (V, E)$, where V is a set of nodes, each node corresponding to one individual in a population of the concerned agent model. E is a set of undirected edges representing contacts between individuals. The prefix *multi* refers to the fact that two nodes can be connected with arbitrary number of edges (see Fig. 2).

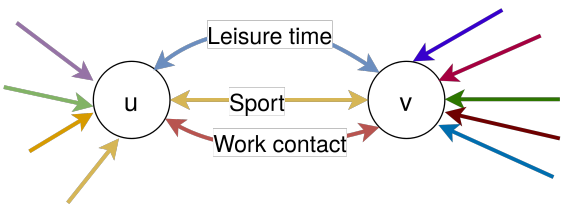

Figure 2: Illustration of two nodes and their edges in a multi-graph.

Edges are divided into 30 layers, each layer refers to a particular type of contact (such as family, work, leisure time, etc.). Layers are associated with weights w that control the activity on the layer as a whole.

Each edge e stores three parameters p_e, i_e and l_e. p_e is the probability that a contact represented by the edge e is realised in the current iteration, i_e is the intensity of that contact, and l_e is the type of the layer.

Each day the edge is activated with the probability

$$p_{active}(e) = w_{l_e} p_e. \tag{1}$$

If the edge is activated and one of its nodes is in an infectious state and the second node is in the state S, the second node is moved to the state E with the probability

$$p_{S\to E}(e) = \beta i_e, \tag{2}$$

where β is a global parameter of the model (however may differ for each node depending on its symptomaticity) and corresponds to the disease modelled.

Policy Module

The Policy Module is an optional, yet important part of the model. It is responsible for a simulation of policies and interventions as will be described in the next section. It communicates both with the model and with the contact graph.

It has access to all parameters of the model and can modify them. In addition it can ask the model to move a particular node to a required state (besides normal state transitions). It has the right to modify the graph, change weights of layers or parameters of individual edges.

It is called between each two iterations of the SEIR model.

3 Simulation of Interventions

The main purpose of the model presented in the previous section is the simulation of various interventions, their study and comparison.

One family of interventions covers protective measures (masks, distancing, hygiene) that make the infection transmission during the contact with an infectious individual less probable. This is modelled simply by reduction of the model parameter β. The parameter β is a vector over all individuals, so individual-based approach is possible. Also asymptomatic individuals have different β than symptomatic ones. In addition, we distinguish between a contact in a family (based on the type of the layer in the graph), where less reduction is applied (no masks in households), and other contacts. The reduction is controlled by the rate parameter ranging from 0.0 to 1.0 that reflects the compliance with protective measures. For the reconstruction of a past progression of the epidemic we use rates based on data from a sociological survey [19].

The other type of interventions covers contact restrictions. The global contact restrictions cover closures of public places, such as school, shops or restaurants. In the context of our model, they refer to restrictions of whole layers. They are modelled by modifying the layer weights w_l (see Eq. (1)). For example, complete closing of schools is realised by setting w_l to zero for all l corresponding to schools. Setting w_l to 0.5 for all l work layers reduces the probability of all work contacts (for example because people switched to home-office). Again, for past epidemic progression reconstruction data from [19] are used to set up the layer weight coefficients.

Individual contact restrictions are the most challenging ones from the modelling point of view. Isolation of individual nodes requires to reduce probabilities of individual edges adjacent to these nodes. This means local temporal modifications of the graph.

If a node v is isolated, we modify the probabilities of all his edges.

$$\forall e \in v_{\leftarrow} : p_e^{new} \leftarrow p_e q_l, \tag{3}$$

where q_l are quarantine coefficients of individual layers and $v_{\leftarrow}$ is a set of edges adjacent to v.

An example of simple quarantine coefficient setup follows:

$$q_l = \begin{cases} 1 & \text{l is a family edge} \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

We implemented three types of policies that use isolation of individual nodes. Namely, they are *self-isolation*, *testing* and *contact tracing*.

Self-isolation simply reflects the fact that a portion of individuals that exhibit symptoms or do not feel well decides to stay home on their own. In the model, as the node starts to exhibit symptoms, it is isolated with a certain probability and it stays in its isolation as long as the symptoms are present.

Second basic policy is the testing. The individuals are tested with a probability θ. If the test is positive, they are marked as detected. We typically use non-zero θ only for symptomatic individuals (for the case of simplicity).

As soon as the individual is detected, it is sent to isolation for a given number of days. The isolation ends with a stop condition, which is optional. It can be two consequent negative tests or just a number of days past from the positive test. As soon as the stop condition is fulfilled, the isolation stops, i.e. the edges of the node are returned to their normal values (unless the second node of the edge is isolated as well).

The last policy is contact tracing. Contact tracing in the real life can be done in an individual manner, algorithmic simulation on the other hand requires a certain level of simplification. The quarantine-detection-isolation life cycle is depicted in the Fig. 4.

As soon as a node is detected it is sent into isolation, i.e. the probabilities of its edges are modified as in (3), the original values are backed up. After a given time, the contact tracing takes place. All edges that were active in the X days before the first symptoms of the node or Y days before its detection are collected. (X, Y are parameters of the contact tracing.)

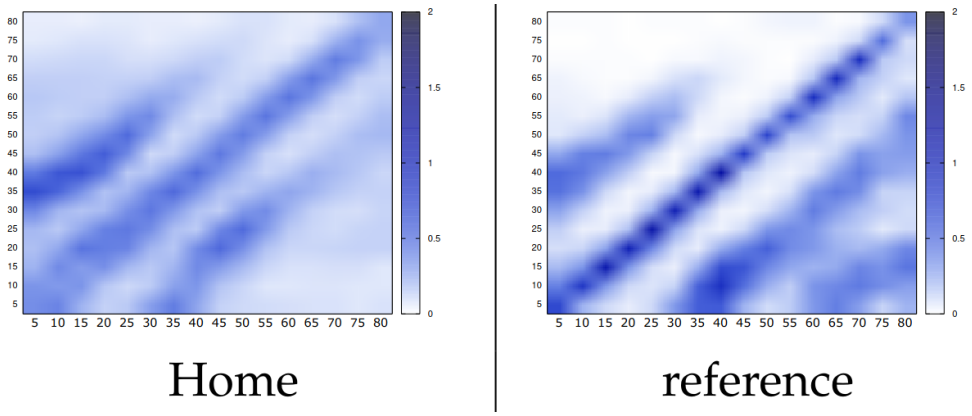

Figure 3: Example of fit of contacts probabilities to Prem matrices. Household contacts in our graph on the left, contacts according to Prem [20] on the right. Axes display age.

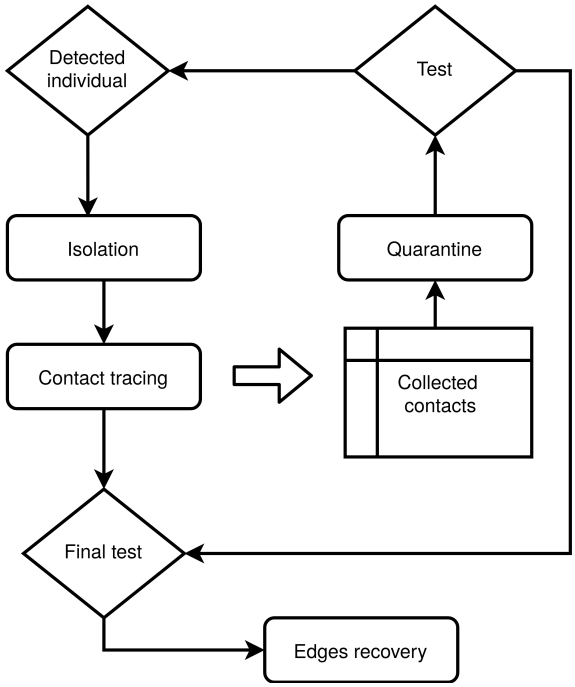

Figure 4: The algorithm of contact tracing procedure.

Then, the edges are filtered — that simulates the fact that in reality all contacts are never successfully traced, some are forgotten, some are not known at all.

The other nodes from the resulting set of edges are sent to quarantine (i.e. isolated). A given number of days after the contact, they are tested. If the test is positive, they undergo the same procedure as detected nodes.

As soon as the node spends a given number of days in isolation (or quarantine), the stop condition is applied in the same way as in the testing policy, and the isolation ends.

The most important parameter characterising the contact tracing is its *strength*. We define it by the tuple (p_1, p_2, p_3, p_4), where the four components correspond to four types of contacts. During the filtering phase, the edge e stays among the collected edges (i.e. is recalled) with the probability

$$p_{recall} = \begin{cases} p_1 & \text{if } e \text{ is a family edge} \\ p_2 & \text{if } e \text{ is a school or work edge} \\ p_3 & \text{if } e \text{ is a leisure time edge} \\ p_4 & \text{otherwise} \end{cases} \quad (5)$$

The policies can be activated or deactivated during the simulation, their parameters may change during the run.

Table 1: Parameters of contact tracing.

parameter	meaning
X	the length of period to collect data before the first symptoms
Y	the length of period to collect data before the detection
C	the delay between detection and collecting contacts
T	the delay between the contact and the test (for contacts)
(p_1, p_2, p_3, p_4)	probabilities of recall (see Eq. (5))
ST	the type of stop criterion

4 Discussion

The model is stochastic, extensively using random numbers, and the variance of results is typically quite high. Therefore in our experiments we always use 1000 simulations for 1000 fixed unique random seeds. Median and mean values are then observed. The time needed for one simulation is approx. 20 seconds on a common CPU.

An example of a possible experiment is depicted in the Fig. 5. There the experiment compares five contact tracing strategies of different strength in the scenario where no other interventions are present.

The strength is defined only with probabilities 1.0 or 0.0 for simplicity, so all contacts from the particular group are always collected. The following strategies were tested: no contact tracing $(0,0,0,0)$, tracing family contacts only $(1,0,0,0)$, tracing family, work and school $(1,1,0,0)$, tracing everything except "others" $(1,1,1,0)$ and ideal tracing $(1,1,1,1)$. We can see that the ideal tracing does not bring much improvement as compared with $(1,1,1,0)$, while the difference between no tracing and tracing family only is more evident.

More experiments on contact tracing strategies comparison and their detailed results can be found in the preprint [7].

The advantage of the model is its modularity. One module can be easily replaced by its alternative. For example, one can replace a contact graph by a contact graph specific to a given environment. Such a graph was a graph built for a school environment, based on real data collected in one Czech school. The results of the simulation study of spread of COVID-19 in a secondary school environment can be found in [9].

5 Conclusion

In the paper we presented a novel agent-based epidemic model. The main difference compared to other models is that it uses a realistic contact graph instead of a synthetic population. Another key feature is the simulation of various non-pharmaceutical interventions that enables to make relative comparisons between them and study their efficacy.

The model is implemented in Python and is publicly available on GitHub [1]. It is configurable and it is possible to write custom policies. An example of a policy we have implemented and not mentioned in this paper is vaccination.

As a future work we plan to include export nodes that will simulate the possibility of infection import from the surrounding world.

Acknowledgement

The work has been supported by the "City for People, Not for Virus" project No. TL04000282 of the Technology Agency of the Czech Republic.

References

[1] Model M, python software, https://github.com/epicity-cz/model-m/releases/.

[2] Openstreetmap. https://www.openstreetmap.org.

[3] State administration of land surveying and cadastre. https://cuzk.cz.

[4] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97, Jan 2002.

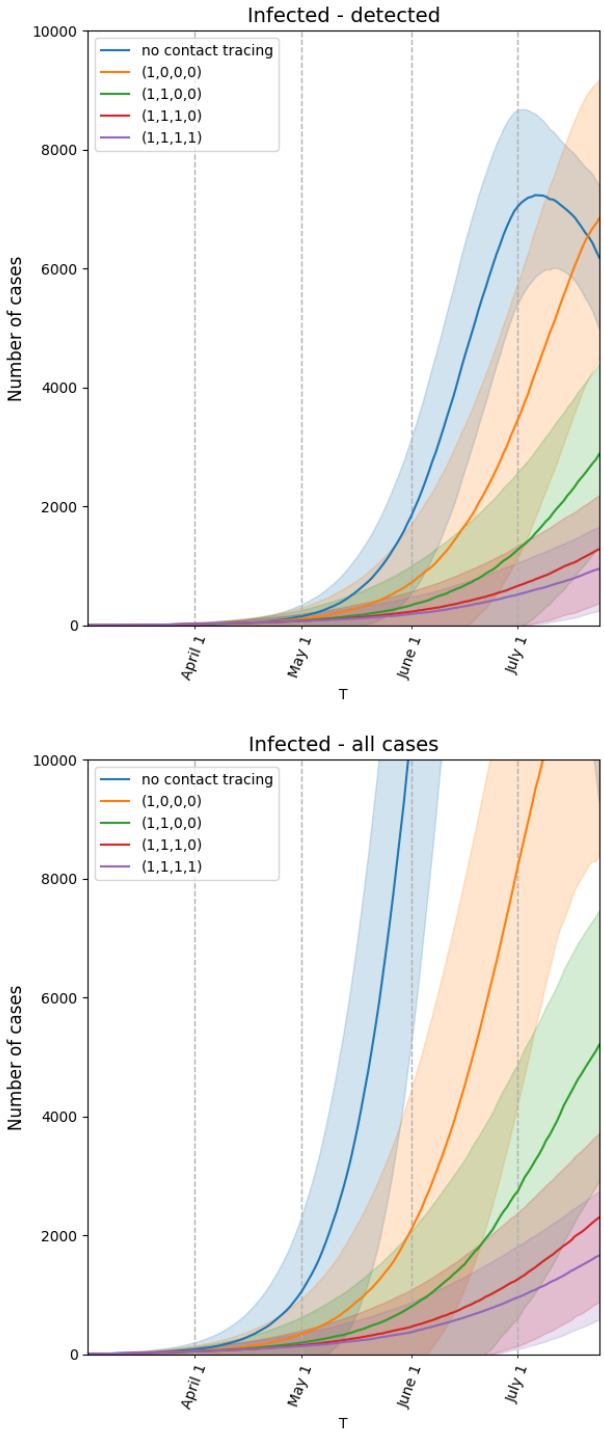

Figure 5: Comparison of five contact tracing strategies. The strength of tracing varies. The median values together with the interquartile range are shown.

[5] Norman TJ Bailey et al. *The mathematical theory of infectious diseases and its applications*. Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE., 1975.

[6] Luděk Berec, Tomáš Diviák, Aleš Kuběna, René Levínský, Roman Neruda, Gabriela Suchopárová, Josef Šlerka, Martin Šmíd, Jan Trnka, Vít Tuček, Petra Vidnerová, Milan Zajíček, and František Zapletal. Model-m: An agent-based epidemic model of a middle-sized municipality. *medRxiv*, 2021.

[7] Luděk Berec, Tomáš Diviák, Aleš Kuběna, René Levínský, Roman Neruda, Gabriela Suchopárová, Josef Šlerka, Martin Šmíd, Jan Trnka, Vít Tuček, Petra Vidnerová, and Milan Zajíček. On the contact tracing for covid-19: A simulation study. 2021. Submitted to The Lancet Public Health.

[8] Alyssa Bilinski, Farzad Mostashari, and Joshua A. Salomon. Modeling Contact Tracing Strategies for COVID-19 in the Context of Relaxed Physical Distancing Measures. *JAMA Network Open*, 3(8):e2019217–e2019217, 08 2020.

[9] Cyril Brom, Tomáš Diviák, Jakub Drbohlav, Václav Korbel, René Levínský, Roman Neruda, Gabriela Suchopárová, Josef Šlerka, Martin Šmíd, Jan Trnka, and Petra Vidnerová. Rotation-based schedules in elementary schools to prevent covid-19 spread: A simulation study. *medRxiv*, 2021.

[10] Czech Statistical Office. Sčítání lidu, domů a bytů 2011. `https://www.czso.cz/csu/czso/scitani-lidu-domu-a-bytu`, 2011.

[11] Matt J Keeling, T Deirdre Hollingsworth, and Jonathan M Read. Efficacy of contact tracing for the containment of the 2019 novel coronavirus (covid-19). *Journal of Epidemiology & Community Health*, 74(10):861–866, 2020.

[12] Cliff C. Kerr, Robyn M. Stuart, Dina Mistry, Romesh G. Abeysuriya, Katherine Rosenfeld, Gregory R. Hart, Rafael C. Núñez, Jamie A. Cohen, Prashanth Selvaraj, Brittany Hagedorn, Lauren George, Michał Jastrzébski, Amanda Izzo, Greer Fowler, Anna Palmer, Dominic Delport, Nick Scott, Sherrie Kelly, Carrie Bennette, Bradley Wagner, Stewart Chang, Assaf P. Oron, Edward Wenger, Jasmina Panovska-Griffiths, Michael Famulare, and Daniel J. Klein. Covasim: an agent-based model of covid-19 dynamics and interventions. *medRxiv*, 2021.

[13] Adam J Kucharski, Petra Klepac, Andrew J K Conlan, Stephen M Kissler, Maria L Tang, Hannah Fry, Julia R Gog, and W John Edmunds. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. *Lancet Infectious Diseases*, 20:1151–1160, 2020.

[14] A. G. McKendrick. Applications of mathematics to medical problems. *Proceedings of the Edinburgh Mathematical Society*, 44:98–130, 1925.

[15] MEDIAN. Market & media & lifestyle –tgi, iii. q. 2019 and iv. q. 2019, 2019. Sociologic survey.

[16] MEDIAN. Interactions and everyday life during covid-19 pandemics., 2020. Sociological survey.

[17] Graham Mooney. "a menace to the public health" — contact tracing and the limits of persuasion. *New England Journal of Medicine*, 383(19):1806–1808, 2020.

[18] MŠMT, Ministry of Eductation Youth and Sports of the Czech Republic. Adresář škol a školských zařízení (list of schools and educational institutions). `http://stistko.uiv.cz/registr/vybskolrn.asp`, 2021. Publicly available database.

[19] PAQ research. Život během pandemie. `https://zivotbehempandemie.cz/`, 2021. Sociological survey.

[20] K. Prem, A. R. Cook, and M. Jit. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Computational Biology*, 13:e1005697, 2017.