Modelling of the Hungarian spread of COVID-19 and control strategies with risk-based approach

Zsuzsa Farkas[1*], Tekla Engelhardt[1], Erika Országh[1], Miklós Süth[1], Szilveszter Csorba[1], Ákos Bernard Jóźwiak[1]

[1]University of Veterinary Medicine Budapest, Digital Food Chain Education, Research, Development and Innovation Institute, H-1078 Budapest, István utca 2., Hungary

*** Correspondence:**
Corresponding Author
farkas.zsuzsa@univet.hu

Abstract

Background

Novel Coronavirus Disease (COVID-19), caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), threatens humanity in terms of health and economy as it spreads extremely fast and causes massive epidemics all over the world. In the absence of a vaccine, social isolation and hygienic measures are the only way to curb the virus.

Methods

In our study, the Hungarian spread of COVID-19 is modelled by applying a modified SEIR (Susceptible, Exposed, Infected, Recovered) compartment model, which takes into account the route of disease transmission not only from infected, but from latent individuals (exposed compartment) as well. The differences between the modified model and the traditional SEIR model has been evaluated. The different scenarios of disease spreading simulate the effect of the different level of interventions (social distancing and hygienic measures) taken place in Hungary. The modelling also considers the population and mobility data which are also essential in case of infectious disease spreading. For controlling the disease in the long-term a network-based analysis is provided based on the concept of the epidemic threshold and the identification of super-spreader population groups.

Results

According to sensitivity analysis of the modified SEIR model, disease transmission of latent individuals has the greatest effect on the number of infections. Based on the results, the applied interventions have a great impact on the disease spreading and are effective in controlling the COVID-19 epidemic., a network-based analysis is provided based on the concept of the epidemic threshold and the identification of super-spreader population groups. According to the results of the network-based study, the proportion of people to be sampled for an effective disease control is the function of the identified people with high number of contacts in social networks who act as super-spreaders.

Conclusion

Applying network-based random, selective and targeted sampling, testing and isolation of affected individuals would yield significantly different sample sizes, highlighting the importance of super-spreaders. Network analysis (but also all computational science methods) need large amount of good quality data and the spread of these methods could be supported by easy-to-use tools. We wanted to raise awareness also on this issue.

Keywords: COVID-19, epidemiological modelling, control strategies, risk-based testing, super-spreader identification, network analysis.

Background

Genetic mutations of microorganisms are inevitable, however, due to climate crisis and other crucial drivers of change, emerging infectious diseases are more and more likely to threaten humanity in the near future. Nowadays, mankind is in a great crisis in terms of health and economy because of an epidemic caused by the mutation of an originally animal-related coronavirus (1)(2). By adapting to humans, the virus named Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) (3) caused pandemic infectious disease COVID-19 (4) that affects millions of lives. As of today, it has appeared almost all over the world, it spreads extremely fast and causes massive epidemics.

COVID-19 is transmitted by inhalation or contact with infected droplets (5,6). It must be noted that the aerosol and surface stability of SARS-CoV-2 is higher than SARS-CoV-1, therefore indirect transmission routes are more significant than previously thought (7,8).

In the absence of a vaccine, social isolation and hygienic measures are the only way to curb the virus. However, these interventions are proved to be effective only when applied very strictly because of the very high virulence of SARS-CoV-2 (9,10). The level of isolation needed (e.g. school closures, banned events and gatherings) is such high that it is unsustainable in the long run.

Modelling is a key option in epidemiology in cases when limited data are available regarding an infectious disease like emerging diseases such as COVID-19. Scenario analysis is a tool by which information can be gained that can support decision making regarding mitigation strategies and risk management, nevertheless, it has its limitations which must be taken into account. As of today, several publications have already appeared which have attempted to simulate the spread of COVID-19 from different aspects. Most of the earliest publications study the epidemic in China by applying the SEIR (Susceptible, Exposed, Infected, Recovered) compartment model with different modifications depending on the purpose of the modelling. Some publications try to estimate the extent of the epidemic regarding time period and number of total infections (11,12), while most of the studies model the effect of interventions such as isolation and quarantine (13–18).

As time goes by, more and more information is available regarding the disease dynamics and disease characteristics of COVID-19, thereby epidemiological models and parameters of the models can more precisely be determined. One of the most important finding is that there are many latent people who develop no symptoms or only mild symptoms, and the disease is infectious in the latent phase as well. Therefore, many people carry and transmit the virus, thereby contribute to the spread of the disease in an unnoticeable way (19–23). Most of the earlier research did not account for this type of transmission because of the lack of knowledge, though it would be essential in understanding the disease dynamics (24).

Conventional compartment models try to capture state transitions and simulate the course of disease in a uniform and homogenous population. However, the basic assumption that any person can contact anyone, and everyone has the same number of contacts, is not true in the real social networks.

Many real networks share same characteristics and are surprisingly similar to each other. These networks are sparse, where many nodes with small number of contacts are connected with each other through few large hubs with many contacts. The number of contacts a node has is denoted by k, and is called the degree of the node. If the degree distribution of a network follows a power law, it is called a *scale-free network* (25*)*. In scale-free networks, compared to random networks, nodes can have very different degrees, ranging from very small, to huge (which are called hubs). Thus, the average degree is not enough in itself to describe the network topology. Diseases can spread much

faster in such networks due to the presence of super-spreader hubs. These networks are also known for their specific robustness: they withstand untargeted attacks without falling apart, but they are vulnerable to attacks targeting the large hubs of the network. This phenomenon could also be used for planning targeted interventions.

The aim of this study is to present an epidemiological modelling based on the SEIR compartment model that is generally accepted for modelling the spread of COVID-19, but with a modification in order to take the transmission route of the latent, yet infectious people into account. Network analysis-based intervention strategies are also discussed in the paper, since the profound implications of network theory are not widely known in the public health community.

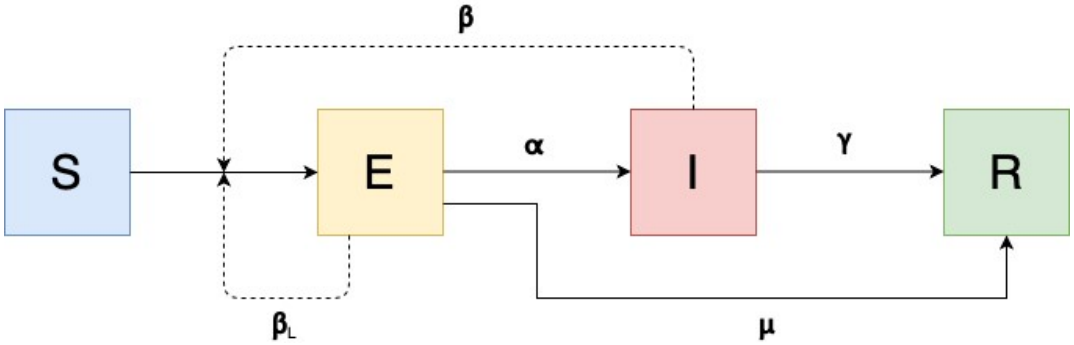
Materials and Methods

Structure of the epidemiological model

Modified SEIR model for COVID-19

For COVID-19, we have built a SEIR (Susceptible, Exposed, Infected, Recovered) compartment model for a basis of the modelling but we have applied some modifications. In the original SEIR model, the Exposed compartment is not infectious. Our modified SEIR model accounts for the infectiousness of this compartment 'E', as it has been shown by previous studies (19–23), and also account for a route of transition of being recovered without symptoms or with mild symptoms from the Exposed compartment to the Recovered compartment. For better differentiation from the original, SEIR model with the applied modifications will be referred as 'modified SEIR model' (Figure 1.), and individuals in Exposed compartment are called 'latent'.

Figure 1



Structure of the modified SEIR model with compartment initials (Susceptible, Exposed (Latent), Infected, Recovered), transitions from one compartment to another (full line), parameters and routes of infection (dashed line). Susceptible individuals can get the infection either from individuals being in the Infected compartment with the rate of β, or from Exposed compartment (latent individuals) with the rate of β_L.

The following differential equations describe the dynamics of the modified SEIR model (1):

$$\frac{dS}{dt} = -\beta_L SE - \beta SI,$$

$$\frac{dE}{dt} = \beta_L SE + \beta SI - \alpha E - \mu E,$$

$$\frac{dI}{dt} = \alpha E - \gamma I$$

$$\frac{dR}{dt} = \gamma I + \mu E \qquad (1)$$

where S, E, I and R are the fractions of Susceptible, Exposed, Infected and Recovered individuals, and S+E+I+R=1 at all times.

Parameters of the model

Selection of values for different model parameters has been done by literature research. There were many estimations in previous publications for the value of basic reproduction number, R_0, ranging from 1.95 (1.4-2.5) (26) to 6.47 (95% CI 5.71–7.23) (27). Our model has been built to be consistent with the average range of reported R_0 values. The mean value of the incubation period was consistently around 4 to 6 days in numerous publications (12,28–31), therefore we selected the number of latent days to be 5 (transition rate from Exposed compartment to Infected compartment). 3.5 days of infectious period was chosen for the calculation of recovery rate, following the studies of Li et al. (19) and Tang et al. (27). The transition rate from Exposed compartment to Recovered compartment arises from the number of days latent individuals spend while carrying and transmitting the virus, which in our model is the sum of incubation period and infectious period, therefor 8.5 days.

The parameter values in the baseline scenario (*Scenario 1*) with no interventions are shown in Table 1.

Table 1. Definitions and values of the modified SEIR model parameters

	Definition	Number of days	Value (day^{-1})
α	Transition rate from Exposed compartment to Infected compartment	5	0.2
γ	Recovery rate	3.5	0.29
μ	Transition rate from Exposed compartment to Recovered compartment	8.5	0.12
β	Transmission rate		0.8*
β_L	Latent transmission rate		0.4*

**Approximate values considered to be in line with the average reported R_0 values (32,33). These parameters will change in Scenarios 2 and 3.*

According to Wu et al. (18) and Nishiura et al. (21), latent individuals are able to transmit the infection 50% less than infected individuals. This is in line with other studies regarding

asymptomatic transmission rate of 50% in case of influenza (34,35). Therefore, the rate β_L (beta latent) will be half of the β value.

Limitations of the compartment model

The formulation was made assuming that natural birth rate is equal to natural mortality rate, therefore the observed dynamics are considered to be caused by the disease. Individual differences, that could influence the susceptibility of people such as age, gender, general health condition cannot be taken into account, all individuals are considered equally susceptible for the disease.

GLEAMviz software

Simulations were performed with the publicly available GLEAMviz Simulation System (36), that is a scientific application designed for performing simulations of the spread of infectious diseases. With the real-world demographic and mobility data, location and time of potential virus transmission by human interactions can be simulated. Compartment model in the Simulation System is customizable, thereby can be adapted for various infectious diseases. The spread of the infection among individuals is driven by the characteristics of the disease specified in the compartment model (37–39).

Setting options regarding compartment models in GLEAMviz

Number of compartments and transitions from one compartment to another (infectious or spontaneous) can arbitrarily be defined. In case of each compartment, air travel and commuting can be allowed or disallowed and there are 3 ways for the settings of infectiousness for each compartment: 1. Carrier 2. Clinical 3. None.

Other setting options in GLEAMviz

Simulations can be done with single run or multiple runs, in latter case, the number of runs can be defined, and the results can be retrieved with median and confidence interval values calculated over the set of runs. Start date and duration of the simulation can be set. GLEAM uses robust statistical methods and executes the simulation in a sequence of time steps that represents full days. Average percentage of airline traffic can be set. For each simulated flight, the number of passengers is a stochastic variable sampled from a binomial distribution whose mean is given by the airline traffic value times the number of bookings. There is an option for enabling seasonality with a built-in algorithm that rescales the basic reproduction ratio R_0 by a sinusoidal function. For commuting model, one can choose from gravity (40) or radiation (41) models. Users can set the average number of hours spent by the commuters at the commuting destination. The default value is 8 hours (the average amount of working time in a day). Minimum number of clinical cases that need to occur in a country for it to be considered infected and minimum number of infected countries for a global epidemic to be considered to occur can be set. Initial population composition for each compartment can be specified in percentages of the total global population. Initial epidemic locations (cities) and number of individuals in the given compartment can be defined for arbitrary number of initial epidemic seeds.

The so-called ‘exceptions’ panel in GLEAMviz provides a way to specify time and space dependent changes of the variable values defined in the compartment model. Each exception lets the user define alternative values for one or more variables for a specified period and specified areas (e.g. cities, countries).

Basic scenarios for modelling disease spread in Hungary with interventions

Scenario 1 – 'worst case'

Worst case scenario with no specific interventions, model parameters determined in the above section 'Parameters of the model'. It represents the conditions of the beginning of the epidemic in Hungary (04 March 2020) (COVID-19 pandemic in Hungary, 2020). It is assumed that only infected individuals with serious and founded COVID-19 symptoms are quarantined in Hungary.

Regarding interventions, our assumption is that all interventions applied (social isolation and hygienic measures) overall decrease the rate of virus transmission, in our case β and β_L values which is in line with the theory of Li et al. (19) and Ferguson et al. (15). Therefore, in case of scenarios with interventions (Scenarios 2, 3), only these parameter values are decreased proportionally depending on the extent of the effect of interventions.

Scenario 2 – First round of interventions

In Hungary, from 16 March 2020, schools and universities operate by distance-learning, nursery schools and other pre-school establishments only provide duty service, more and more employers provide the opportunity of home office for the workers and people are getting conscious about staying at home and reducing the physical contacts as much as possible. Events with large number of people were banned (the limit was 100 people for indoor and 500 for outdoor events, respectively). According to our assumption that is supported by the data of Google Analytics (42) and Hungarian traffic statistics (43), that means a 50% decrease in the level of social interactions compared to the normal lifestyle. This, together with the raised awareness of the disease and the importance of detection and quarantine of infected people, based on the approach of Ferguson et al. (15) results in the decrease of β and βL values to 0.4 and 0.2 (from 0.8 and 0.4), respectively.

Scenario 3 – Second round of interventions

From 28 March, the Hungarian government have ordered curfew restrictions across the country, in the first round until the 11th of April, and in the second round it was extended for an indefinite period. The goal of the restrictions was generally to limit the contact of people who do not live in the same household, therefore residences should only be left for the satisfaction of basic needs (e.g. work, grocery, pharmacy, health services) the previously specified restrictions regarding the institutions have remained unchanged. According to our assumption and the further mentioned references (42,43) that resulted in a further 25%, altogether 75% reduction in the level of social interactions compared to the normal lifestyle, which means a reduction in β and β_L values to 0.3 and 0.1, respectively.

Scenarios 1 to 3 attempt to model the real-life situation regarding disease spreading with the temporally applied intervention measures in Hungary. It is noted, however, that the list of interventions is not fully complete, only those are mentioned which are crucial regarding scenario building.

Practical implementation in GLEAMviz (Settings)

Unless not specifically indicated, settings are applied for each four Scenarios.

'MODEL' panel: Modified SEIR model (Figure 1.) with parameters defined in section 'Parameters of the model'. Special GLEAMviz setting options for different compartments are shown in Table 2.

Table 2. Special GLEAMviz setting options in different compartments

	Susceptible	Exposed (Latent)	Infected	Recovered
Infectiousness	none	carrier*	clinical**	none
Commuting	allowed	allowed	disallowed	allowed
Air travel	allowed	allowed	disallowed	allowed

*: *but transmitting the virus with the rate of* β_L

**: *transmitting the virus with the rate of* β

'SETTINGS' panel:

- Multi-run simulation (20 runs, the maximum possible number of runs)
- Start date: 4 March 2020 - the date of the first reported cases of SARS-CoV-2 infection (44)
- Airline traffic:
  - *Scenario 1*: 88% (12% decrease globally in the first half of March according to OAG statistics) (45)
  - *Scenario 2*: 52% (48% decrease globally in the second half of March) (45)
  - *Scenario 3*: 30% (66% decrease rounded up to 70% as data were available only until 20 April 2020)(45)
- Commuting model: radiation model was chosen as it is proved to describe the commuting patterns better when parameter-free algorithms must be used (41,46).
- Time spent at commuting destination:
  - *Scenario 1*: 8 hours (default, assuming no change at commuting in Hungarian region)
  - *Scenario 2*: 4 hours (50% reduction in commuting according to Hungarian traffic statistic)(42,43)
  - *Scenario 3*: time spent at commuting destination: 2 hours (75% reduction in commuting according to Hungarian traffic statistic) (42,43)
- Seasonality: disabled – as it is believed that SARS-CoV-2 is not affected by weather parameters and there is no evidence to the contrary at the moment (47).
- Minimum number of clinical cases that need to occur in a country for it to be considered infected: 1 (default)
- Minimum number of infected countries for an occurrence to be an epidemic: 2 (default)
- Initial global distribution of a population in compartments: 100% susceptible
- Initial geographic location of the epidemic: Officially 2 infected people were registered on the 4th of March 2020 in Hungary (48), Budapest, but in order to get more realistic simulations regarding the coverage of the region, initially 5 infected individuals were set for Budapest and 3, 3 latent individuals for Debrecen and lake Balaton, based on population data, respectively. These were the only available locations in GLEAMviz software for Hungary.
  As there is a possibility for setting numerous initial locations of the epidemic, number of individuals with registered COVID-19 infections (number of active patients) were collected for the 4th of March, 2020. Only cities can be set in the software, and data about registered

infections were available in different levels of aggregation. Table 3. shows the initial locations with number of infected people used in the simulation (only >20) apart from Hungary.

Table 3. Initial locations and number of infected individuals (apart from Hungary) set in GLEAMviz.

Initial location		No. of infected individuals	Level of aggregation of the data	Reference
Country	**City**			
Austria	Innsbruck	22	national (halved)	(49)
	Vienna	21	national (halved)	
France	Lille	65	regional	(50)
	Lyon	49	regional	
	Strasbourg	38	regional	
	Paris	55	regional	
Germany	Düsseldorf	115	federal state	(51)
	Munich	48	federal state	
	Stuttgart	50	federal state	
Italy	Ancona	80	regional	(52)
	Bologna	516	regional	
	Florence	37	regional	
	Milan	1497	regional	
	Naples	31	regional	
	Rome	27	regional	
	Turin	82	regional	
	Venice	345	regional	
Spain	Barcelona	24	autonomous community	(53)
	Madrid	90	autonomous community	
China	Beijing	1267	national except Hubei	(54)
	Wuhan	24085	Hubei province	
Iran	Tehran	2259	national	(54)
Japan	Tokyo	278	national	(55)
Singapore	Singapore	33	national	(56)
South Korea	Seoul	5547	national	(57)

Settings of 'EXCEPTIONS' panel are shown in Table 4.

Table 4. 'EXCEPTIONS' panel settings in GLEAMviz.

No. of Scenario	Exception(s) applied	
Scenario 1	No exceptions applied	
Scenario 2	$\beta = 0.4$; $\beta_L = 0.2$; from 16 March 2020 till the end of the simulation; in the Hungarian region	
Scenario 3	$\beta = 0.4$; $\beta_L = 0.2$; from 16 March 2020 to 27 March 2020; in the Hungarian region	$\beta = 0.3$; $\beta_L = 0.1$; from 28 March 2020 till the end of the simulation; in the Hungarian region

Comparison of modified SEIR model with traditional SEIR model

In order to evaluate the differences between modified SEIR model and the traditional ('basic') SEIR model, a simulation has been made in which parameters β_L and μ were disregarded (transition rates that contribute to the infectiousness of exposed individuals). Other parameters and settings were the same as in Scenario 1 in the simulation.

Sensitivity analysis of modified SEIR model

The sensitivity analysis has been done to reveal the parameter(s) to which the model is the most sensitive. There were five parameters in our modified SEIR model to be investigated: β, β_L, γ, α, μ. The parameters and settings of *Scenario 1* was used for the evaluation. The maximum value of the daily number of individuals in the infected compartment and the number of days related to this value were selected as the endpoint. Only one parameter per scenario was changed at a time. Changes in both directions were evaluated with lower ($\times$ ½) and higher ($\times$ 2) values compared to the baseline value (Table 1.) of the examined parameter.

Network analysis-based intervention strategies

The basic assumption of the conventional compartment models is that any person can contact anyone, and everyone has the same number of contacts, is not true in the real contact networks. Real networks are sparse, where many nodes with small number of contacts are connected with each other through few large hubs with many contacts, which could also be identified as super-spreaders.

The first applications of network science to disease modelling set a new scientific field called network epidemics (58). When modelling disease spreading, the network characteristics will yield many important differences compared to the conventional compartment models.

First of all, the diseases spread much faster in scale free networks due to the presence of hubs. In network epidemics the concept of *epidemic threshold* (λ_c) is used: pathogens can only spread if the spreading rate λ exceeds λ_c. The spreading rate can be defined as:

$$\lambda = \frac{\beta^*}{\mu} \qquad (2)$$

where β^* is the likelihood that the disease will be transmitted from one infected person to a susceptible one in a unit of time, and μ is the recovery rate. The conventional R_0 can be defined as:

$$R_0 = \langle k \rangle \lambda \qquad (3)$$

where $\langle k \rangle$ is the average degree. In case of scale free networks, the epidemic threshold is:

$$\lambda_c = \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1} \qquad (4)$$

where $\langle k^2 \rangle$ is the second moment of the degree distribution, and it is used in the calculation of the variance (25). In scale free networks, $\langle k^2 \rangle$ is significantly larger than $\langle k \rangle$, due to the huge variance in degrees of the nodes. This means, that in large scale-free networks λ_c is very small, meaning that most of the diseases can spread very rapidly because of the presence of large hubs. This holds for every network (not only scale-free ones), where there is a large difference between the degrees of the nodes. This phenomenon was also captured by Meyers et al. (59), who applied the methods of contact network epidemiology for SARS, to illustrate that for a single value of R_0, any two outbreaks, even in the same setting, may have very different epidemiological outcomes.

Many social contact networks are also scale-free networks. We don't have a map of the Hungarian contact network of people, but we can estimate λ_c with the use of other data. The Copenhagen Networks Study resulted in a multi-layer temporal network which connects a population of more than 700 university students over a period of four weeks (60). We have used a 2-week (common agreed suggested quarantine period)(61) subset of the network of physical proximity among the participants (estimated via Bluetooth signal strength) to simulate direct personal contacts. We have used the settings defined by Stopczynski et al. (62): only contacts where the signal strength was ≥ -75 dBm were used, corresponding to distances of approximately 1 meter or less. Based on these data, a network had been constructed and analysed with the use of KNIME free and open-source data analytics (63), reporting and integration platform and igraph R-package (64).

Other important consequence of network topology is the difference in planning intervention strategies. If we try to remove nodes from the network either with immunization (which we can't perform yet in case of COVID-19) or with identifying latent and infected people with sampling and testing, and then removing them with quarantine measures, we may have use other than conventional options as well. For assessing the effect of identifying and removing latent and infected people with sampling, testing and consequent quarantine measures, we use the concept of critical immunization well established in network epidemiology. Critical immunization g_c is the proportion of nodes needed to be removed from the network to stop the spreading of the disease.

Random sampling

In case of random sampling, testing (and consequent removing of SARS-CoV-2 positive people from the contact network), the critical immunization g_c can be defined as:

$$g_c = 1 - \frac{\lambda_c}{\lambda} \qquad (5)$$

Besides random sampling, there are other options as well, stemming from the characteristic of the scale free networks first described as the error tolerance of networks (65). If we remove the nodes from a network in a targeted manner, the spreading on the network could be quickly slowed down or

stopped. As it was demonstrated, the rapid spread of diseases is caused by the fact that there is a large variance in the degrees of the nodes in these networks ($\langle k^2 \rangle \gg \langle k \rangle$), and this comes from the presence of hubs. If we want to decrease the variance (thus increasing λ_c), we have to block hubs from interacting.

Selective sampling

If we don't know the exact mapping of the contact network, we could reach for the 'friendship paradox' (66) and the immunization strategy based on it proposed by Cohen et al. (67). The friendship paradox says that on average the neighbours of a node have higher degrees than the node itself. The average degree of a node's neighbour doesn't equal to $\langle k \rangle$, but it is a different number, depending largely also on $\langle k^2 \rangle$ (25). The origin of this phenomenon is that it is more likely for a random node to be connected to a hub than to a small degree node, because hubs have more connections than other nodes. Thus, immunizing (or isolating) the contacts of randomly selected individuals, we target the hubs without knowing exactly which individuals are the hubs.

Targeted sampling

If we knew the whole contact network, we could target the most prominent hubs. We don't know the exact mapping of the Hungarian social network, but we can have the following assumptions:

- The network is scale-free;
- where the probability p_k that a node has exactly k links is: $p_k = \frac{k^{-\gamma}}{\xi(\gamma)}$, where γ is the degree exponent and $\xi(\gamma)$ is the Riemann-zeta function (using discrete formalism (25));
- and $2 < \gamma < 3$, as in most of real-life networks.

With targeted sampling we try to remove all nodes whose degree is larger than k_t. From an epidemiological viewpoint this is the same as removing the high degree nodes from the network with their links as well. With this intervention, the network will change, and λ_c will increase (25):

$$\lambda_c' = \left[\frac{\gamma-2}{3-\gamma} k_t^{3-\gamma} k_{min}^{\gamma-2} - \frac{\gamma-2}{3-\gamma} k_t^{5-2\gamma} k_{min}^{2\gamma-4} + k_t^{2-\gamma} k_{min}^{\gamma-2} - 1\right]^{-1} \qquad (6)$$

Depending on the degree exponent, and setting k_{min} to 1, we could obtain target degrees (k_t) with the new epidemic thresholds (λ'_c).

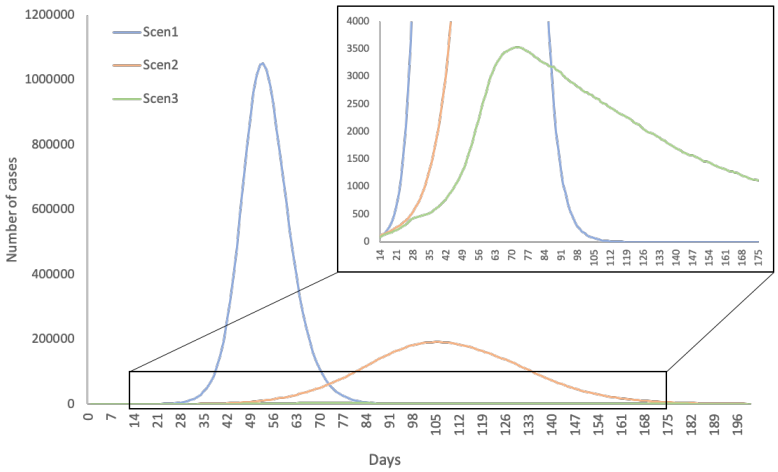
Results

Results of epidemiological modelling

Different scenarios were run in the publicly available version of GLEAMviz software with the settings described in 'Methods' section. As the software only provides the number and cumulative number of new transitions of individuals per 1000 people regarding one compartment to another in a daily breakdown, an algorithm was developed in KNIME software (63) in order to automatically convert the GLEAMviz outputs into actual daily case numbers of different compartments.

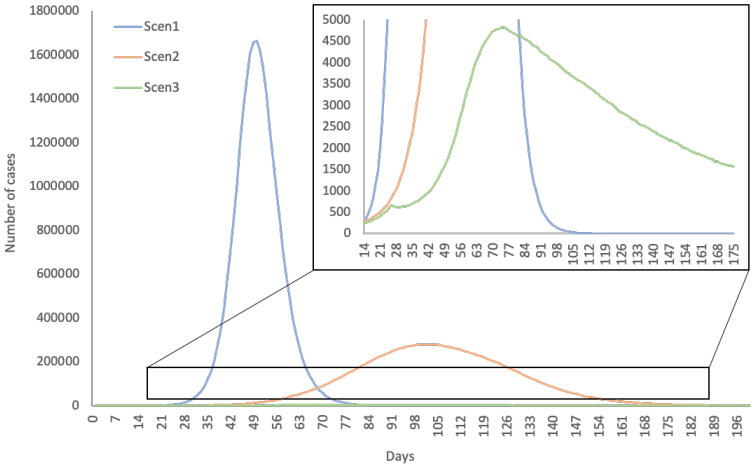
Figures 2 and 3 show the comparison of distribution of Infected and Exposed (latent) individuals in different scenarios.

Figure 2



Distribution of Infected individuals in *Scenarios 1* to *3*

Figure 3



Distribution of Exposed (latent) individuals in *Scenarios 1* to *3.*

The proportion of maximally affected latent and infected people (daily and overall) can be seen in Table 5. According to the modelling, about 1,000,000; 190,000 and 3,500 people will be infected on the day, when the epidemic reaches its peak in *Scenarios 1, 2* and *3,* respectively. Overall, about 5,400,000; 3,300,000 and 110,000 people will get through the infection with moderate or serious symptoms in *Scenarios 1, 2 and 3* and respectively. Note that in case of Scenario 3, the maximum number was not reached till the end of the simulation (365 days, which was the limitation of the software).

Table 5. Proportion of maximum number of daily and cumulated cases of infected and latent people in Hungary.

			Proportion in the population[1] (median with LCL[2] and UCL[3]) (%)	No. of days[4]
Scenario 1	Maximum of daily cases	Infected	11.42 (7.99-12.62)	52
		Latent	18.07 (12.51-20.29)	49
	Maximum of cumulated cases	Infected	58.69 (58.65-58.72)	118
		Latent	93.89 (93.85-93.91)	122
Scenario 2	Maximum of daily cases	Infected	2.08 (1.76-2.38)	105
		Latent	3.05 (2.57-3.05)	101
	Maximum of cumulated cases	Infected	35.89 (35.82-35.94)	268
		Latent	57.36 (57.26-57.42)	279
Scenario 3	Maximum of daily cases	Infected	0.04 (0.04-0.04)	72
		Latent	0.05 (0.05-0.06)	74
	Maximum of cumulated cases	Infected	1.22 (1.12-1.31)	365*
		Latent	1.85 (1.69-2.00)	365*

as 9,213,366 people
2: Lower confidence limit originating from the 20 different simulations
3: Upper confidence limit originating from the 20 different simulations
4: Regarding median value
*: Maximum number was not reached till the end of the simulation (365 days, which was the limitation of the software)

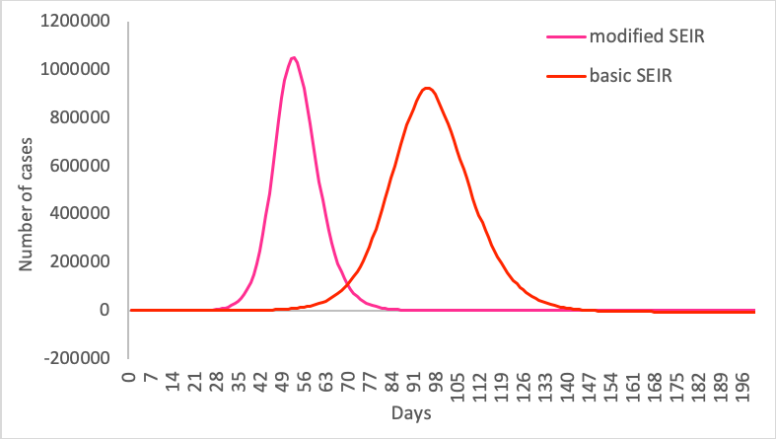
Regarding exposed (latent) people, who will only have mild symptoms or no symptoms at all (according to our assumption), on the day of the peak of the epidemic, about 1,700,000; 280,000 and 4,800 people will be affected in *Scenarios 1, 2 and 3*, respectively. Overall, about 8,700,000; 5,300,000 and 170,000 people will get through the infection with mild symptoms or as asymptomatic cases in *Scenarios 1, 2 and 3,* respectively. The same limitation in case of Scenario 3 applies here as well regarding the maximum number of latent people.

Note that the modelling is a robust estimate for actual case numbers, therefore only approximate values are indicated.

Comparison of modified SEIR model with traditional SEIR model

Basic SEIR model results in somewhat lower number of daily infections, daily latent cases, and cumulative latent cases (1.38%, 3.07%, and 1.74%, respectively) but overall, the cumulative number of infectious individuals are 33.47% higher compared to the modified SEIR model (Table 6). It can be seen that the peak of the curve is shifted in the basic SEIR model (from about 50 days after the start date of the simulation to the ~90th day) (Fig. 4.).

Figure 4



Comparison of the distribution of Infected individuals in case of modified SEIR model and basic SEIR model in case of *Scenario 1*

Table 6. Proportion of maximum number of daily and cumulated cases of infected and latent people in modified SEIR model compared with basic SEIR model Settings of Scenario 1 was applied.

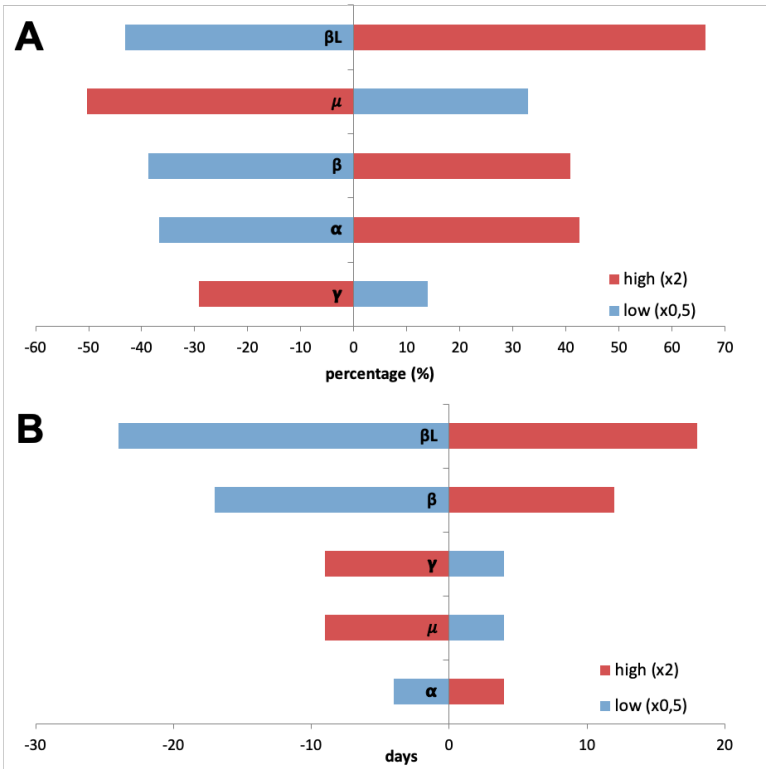
		Modified SEIR model		Basic SEIR model	
		Proportion in the population[1] (median with LCL[2] and UCL[3]) (%)	No. of days[4]	Proportion in the population[1] (median with LCL[2] and UCL[3]) (%)	No. of days[4]
Maximum no. of daily cases	Infected	11.42 (7.99-12.62)	52	10.04 (8.01-10.89)	95
	Latent	18.07 (12.51-20.29)	49	15.00 (11.61-16.68)	91
Maximum no. of cumulated cases	Infected	58.69 (58.65-58.72)	118	92.16 (92.10-92.22)	214
	Latent	93.89 (93.85-93.91)	122	92.15 (92.10-92.20)	211

[1]: Population of Hungary registered in GLEAMviz was 9,213,366 people
[2]: Lower confidence limit originating from the 20 different simulations
[3]: Upper confidence limit originating from the 20 different simulations
[4]: Regarding median value

Sensitivity analysis of modified SEIR model

Most pronounced changes in both chosen endpoints can be seen in case of β_L, increasing its value by 2 times resulted in a ~66% increase in the number of maximum daily infections (Fig. 5A), and decreasing its value to the half of the original ended up in a 24 days sooner peak in the curve of the number of daily infections (Fig. 5B). However, other the importance of parameters such as μ and β are also indicated by the tornado-plot.

Figure 5



Results of sensitivity analysis of the modified SEIR model: changes in daily maximum value of individuals in the infected compartment (A) and changes in the number of the days related to the former endpoint (B)

Results of network analysis

To address the problem caused by the inability of conventional compartment models to capture the non-homogenous nature of the population, network-based analyses were also performed, which yielded the following results.

Based on the Copenhagen Networks Study (60) data, we found average degree $\langle k \rangle$=46, the second moment of the degree $\langle k^2 \rangle$=2847 and based on Equation (4) the epidemic threshold λ_c = 0.016. Basic parameters of modified SEIR model (*Scenario 1*) was set to be in line with the average R_0 value of 3.1 of Read et al. (32) and Tian et al. (33). Using Equation (3), the spreading rate of the virus was λ=0.067, which means that for stopping the epidemic, we need to change the network in such a way that λ_c will be increased above this value.

Random sampling

In case of random sampling, testing (and consequent removing of SARS-CoV-2 positive people from the contact network), based on equation (5) we have found the critical immunization to be g_c=0.761. This means that 76.1% of the population shall be removed from the network. This implies very strict sampling and quarantine measures with testing a very large proportion of the population.

Selective sampling

Using the 'friendship paradox', the procedure proposed by Cohen et al. (67), consists of three steps:

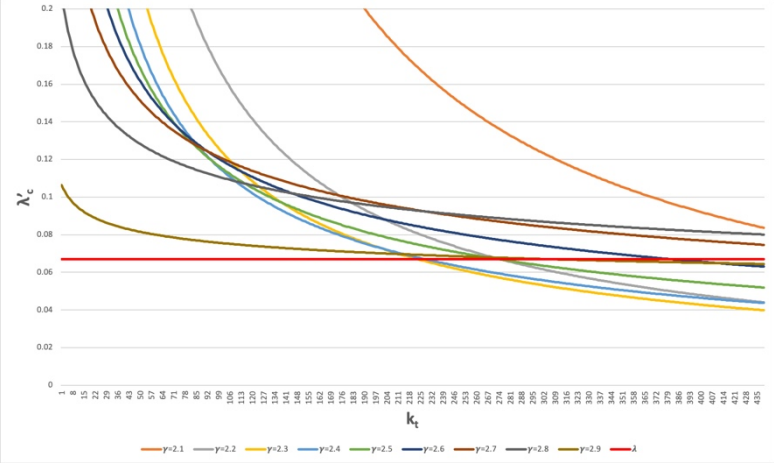
1. Choosing randomly a fraction of nodes, this is layer 0.
2. Selecting randomly a link for each node in layer 0. The nodes to which these links connect will form layer 1.
3. Immunizing (in our case sampling and testing) the layer 1 individuals.

This sampling strategy doesn't require information on the structure of the network. According to Cohen et al. (67), g_c is systematically under 0.3. It means that by selecting and removing a randomly chosen neighbour of 30% of the population, the spreading of the disease could be stopped. With selecting contacts referred simultaneously by more people, this strategy could be even enhanced.

Targeted sampling

With targeted sampling we try to remove all nodes whose degree is larger than k_t. Based on equation (6), we could obtain target degrees (k_t) with the new epidemic thresholds (λ'_c) which are presented on Figure 6. The figure shows target degrees k_t in the range of 234 (for γ=2.3) *to >1000* (for γ=2.8), meaning that if we could sample (and then contain all positive cases) the super-spreaders with contact number above 234, the epidemic threshold would be high enough to stop the disease.

Figure 6



Epidemic thresholds (λ'_c) of different degree exponents (γ) and target degrees (k_t) calculated to increase the epidemic threshold above the spreading rate of COVID-19 with targeted sampling. The spreading rate (λ) of the virus is set to 0.067 (red line).

How does it translate to the Hungarian population? Idealizing a scale-free network, the number of people with contacts $k \geq k_t$ can be calculated with (25):

$$N_{k \geq k_t} = N \sum_{k=k_t}^{\infty} \frac{k^{-\gamma}}{\xi(\gamma)} \qquad (8)$$

Setting γ=2.3 and k_t=234, this is in the magnitude of 0.04% of the population only.

Discussion

Potentials and limitations of the modelling

Scenarios have been designed to simulate the most likely spread of COVID-19 in Hungary with no interventions (*Scenario 1*) and the effect of actually applied interventions *(Scenarios 2 and 3)* in order to gain information on the effect of intervention measures on the disease spreading. According to the results, the applied interventions (social distancing and hygienic measures) have a great impact on the disease spreading and are effective in controlling the COVID-19 epidemic (*Scenario 3*).

The main advantage of epidemiological modelling with GLEAMviz software is that besides the characteristic parameters of the infectious disease, other essential factors are taken into account when simulating the disease spreading, namely population density and mobility of the population. The irregular network structure, that affects the local spread of infectious disease between neighbouring subpopulations is captured in GLEAMviz datasets, so as the difference between high traffic and low traffic airports that has a significant impact of disease spreading around the globe.

Simulations are evaluated for the Hungarian region, but with GLEAMviz, it is possible to set the initial number of infected individuals in different cities on the day when the simulation starts, thereby the impact of global presence of infected people is also taken into account. Note that registered number of infected people in different countries/cities carries bias as the protocol for registering COVID-19 positive cases and number of tests performed differ greatly from country to country. Nonetheless, these data and this option provide additional adjustable settings that contributes to more realistic simulations.

Evaluation of modified SEIR model – sensitivity analysis

All modelling has their limitations, and this applies particularly to modelling of emerging diseases such as COVID-19. As limited data are available regarding SARS-CoV-2 and the spread of COVID-19, parameters of the compartment model can only be estimated. Our modified SEIR model has been made by using available data from scientific literature regarding the compartment model parameters and its novelty is that the route of infection transmission from the Exposed compartment (latent individuals) is built in. Compared to the traditionally applied SEIR model, in which latent individuals cannot transmit the infection to susceptible people, it results in an earlier and higher peak in the epidemic curve of infected individuals, which means a faster epidemic course with more infected people in shorter period of time, that poses greater burden to the healthcare system. However, total number of infections are lower altogether in the modified model compared to the basic SEIR model. According to this result, the experienced characteristics of COVID-19 epidemics, namely the so-called ‘exponential growth’ of the disease spread can be explained by the disease transmission of latent people in case of SARS-CoV-2. Sensitivity analysis of the modified SEIR model also reveals the importance of the disease transmission by latent individuals. According to the results, disease transmission related to latent individuals has the greatest impact on the results of the simulation regarding the number of maximum daily infections and the time of the peak of the epidemic.

Evaluation of network analysis results

When taking into account the social contact network topology during intervention planning, we have to be aware of the specific challenges and also opportunities these network characteristics pose. Most of the counterintuitive network phenomena are caused by the presence of hubs, i.e. nodes with large amount of contacts, which could also be identified as super-spreaders. One of the results is that the diseases spread much faster on real networks than conventional compartment models would predict. This also means that the R_0 identified as a benchmark value might be wrong: even if $R_0 < 1$, there

might be still nodes in the network whose degree is higher than the average degree, maintaining and spreading the disease.

The usual intervention strategies mentioned in this paper are still useful and effective, but they don't target the super-spreader hubs. When planning sampling strategies, *random sampling* would imply very strict sampling and quarantine measures with testing a very large proportion of the population (>75%). When the 'friendship paradox' (i.e. on average the neighbours of a node have higher degree than the node itself) would be used for planning *selective sampling*, a significantly lower number of samples (~30%) would suffice. And if we could find the super-spreader hubs in a network during *targeted sampling*, <1% of the population would be enough to be sampled, tested and isolated.

Of course, natural contact networks are not ideal scale-free networks, and we don't know the exact super-spreaders either. Nevertheless, according to the recent analysis of the Hungarian Centre for Economic and Regional Studies Institute of Economics on occupations affected the most by the COVID-19 outbreak (69) 18% of the people are working in these occupational areas. These include healthcare and social workers, retail and restaurant workers, drivers, cleaning staff, pharmacists, veterinarians, public transport workers, postmen and waste removal staff. If the targeted sampling strategy would aim for these occupations, with strict quarantine measures of the positive cases, the spreading of the virus could be significantly slowed down.

Conclusion

Comparison with real COVID-19 epidemiological situation in Hungary

The first wave of the disease spreading, according to the reported COVID-19 cases in Hungary, SARS-CoV-2 have not caused a great epidemic in Hungary Multiple reasons can explain this, some of them are listed hereunder.

- By the time the epidemic reached Hungary, there were several frightening scenarios (massive amount of infections and deaths caused by COVID-19 in China, Italy etc.) that warned the public and raised the awareness for caution about COVID-19 and discipline regarding social distancing and hygiene. People have started to draw each other's attention through social media and encouraged staying at home, therefore social distancing have started even earlier and have been applied stricter than the actual measurements of the Hungarian government required.
- Because of conscious social distancing, super-spreading events have been prevented in the very early phase of the epidemic, which has a great impact on the fate of later number of infections (59).
- Many studies investigate the association of Bacillus Calmette-Guérin (BCG) vaccination with reduced morbidity and mortality of COVID-19 (70,71). In Hungary, BCG vaccination is compulsory since 1954, therefore BCG vaccination coverage is very high. This can also contribute to the low number of COVID-19 infections in Hungary.

Our modelling results show a good correlation with the results of the Hungarian CoronaVirus disease-19 Epidemiological Research(72). H-UNCOVER was a large-scale, cross-sectional, representative population screening of Hungarian male or female patients aged 14 and over who live in private households, aiming for sampling and investigating the rate of the infected patients, asymptomatic carrier and healed patients (48). According of the extrapolation of the results of 10,575

patients sampled between 1-16 May 2020, there were 2,421 active cases in Hungary and 56,439 people were asymptomatic carriers and healed patients.

According to the results of our modelling in *Scenario 3*, there were 3,527 active infected cases on day 70 (12th May) and 38,692 cumulative recovered people by day 70 and 4,756 active latent cases on day 70, adding up to 43,448 asymptomatic carriers and healed patients.

Proposal for disease control with agile, risk-based testing

Our study proposes a long-term feasible risk-based approach for testing SARS-CoV-2 infections, that could gradually replace isolation interventions by the early detection of positive cases between the so-called 'super-spreaders', thereby preventing mass infections and breaking infection chains with removing the seed of the infection in a timely manner. In cases when number of positive cases tend to increase nonetheless of the quarantine of affected people, social isolation measures can be re-initiated in a timely manner and epidemic crisis can thereby be prevented.

With the help of network analysis, people can be grouped into risk categories regarding SARS-CoV-2 transmission. In this case, the key parameter for establishing risk categories is the number of physical interactions between people. A continuous testing is suggested but with the allocation of resources (test kits, testing personnel, financial resources) to the high-risk groups. People could be categorized by risk scores of physical interactions based on their occupation. Occupations like healthcare and social workers, retail and restaurant workers, drivers, cleaning staff, pharmacists, veterinarians, public transport workers, postmen and waste removal staff would fall into this high-risk group. These people must be tested much more often as they play key role in SARS-CoV-2 transmission. From this group, healthcare workers could be identified as key actors from network perspective, implying even stricter sampling and testing regime for them. When identified positive cases of super-spreaders would be isolated, this would have a large effect on the contact network itself, cutting off all the links of the large hubs, thus slowing down the spreading considerably.

Our approach is generally applicable for the prevention and early detection of epidemics caused by other microorganisms as well, thereby protecting human health and preventing economic crises caused by emerging and re-emerging diseases.

Proposal for data collection and data sharing

Our study also points to the well-known fact that models are only as good as their input data. When modelling and planning intervention strategies, fit-for-purpose and timely data are essential. Using specific population like college student data can't be used for precise modelling of the spreading of diseases in other societal groups. Unfortunately, there is a lack of social contact data with sufficient granularity, and advances on network science couldn't be fully exploited in real-life situations. It is not to say that collecting and sharing contact network data for public health purposes would overwrite data protection and privacy aspects, but as Oliver et al. (73) pointed out, there are available data sources like mobile phone data, which could be extremely useful for such purposes. These datasets, if would be made available in a careful and transparent manner and taking into account data protections issues, could also be used for other important public health domains, like foodborne disease outbreak investigation.

In our paper we had two important objectives regarding network analysis approach to epidemiological modelling:

- Showing the profound implications of network theory for epidemiological modelling and risk management, since the network epidemiological approach is not widely known in the public health community. Most of the analyses and also intervention strategies don't take into account the inhomogeneity of the connections nor the network-based background of the spreading of the virus. Even when there is a general knowledge on the role of super-spreaders, the quantification of this role is not well known.

- Network analysis (but also all computational science methods) need large amount of good quality data and the spread of these methods could be supported by easy-to-use tools. We wanted to raise awareness also on this issue.

These two objectives are in close connection, since the lack of fit-for-purpose data and access to computational methods prevent the public health community from applying network-based approach in the decision-making process. However, the public health community should work towards solving data and tool related issues, for example with using proxy data in a short term (e.g. mobile phone data, tracing applications, etc.) or with planned exercises on network data collection in the long run.

Declaration

Ethics approval and consent to participate
Not applicable.

Consent for publication
The authors consent for publication.

Availability of data and materials
All the analysed and generated data during this study are included and referred in this published article. The KNIME workflow is available at https://univet.hu/en/research/dfi/publications/scientific-publications/covid-19/.

Competing interests
The authors declare no competing interests.

Funding
Not applicable.

Author Contributions
ZSF prepared the modelling and wrote the main body of the text; ÁJ created the concept of the network study and prepared the related research; TE and ZSF prepared the tables; SZCS prepared the figures; MS, EO, SZCS, TE and ZSF did the literature research; MS reviewed the first draft of the manuscript. TE prepared the sensitivity analysis. All authors contributed to the creation of the epidemiological compartment model in the study and the concept of the research. All authors contributed to manuscript revision, read and approved the version to be submitted

Acknowledgement
Not applicable.

References

1. Zhou P, Yang X-L, Wang X-G, Hu B, Z. hang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020 Mar 12;579(7798):270–3.

2. Zheng J. SARS-CoV-2: an Emerging Coronavirus that Causes a Global Threat. Int J Biol Sci. 2020;16(10):1678–85.

3. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol. 2020 Apr;5(4):536–44.

4. WHO. Naming the coronavirus disease (COVID-19) and the virus that causes it [Internet]. 2020. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it

5. Singhal T. A Review of Coronavirus Disease-2019 (COVID-19). Indian J Pediatr. 2020 Apr;87(4):281–6.

6. Cascella M, Rajnik M, Cuomo A, Dulebohn SC, Di Napoli R. Features, Evaluation, and Treatment of Coronavirus. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 [cited 2020 Nov 25]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK554776/

7. van Doremalen N, Bushmaker T, Morris DH, Holbrook MG, Gamble A, Williamson BN, et al. Aerosol and surface stability of HCoV-19 (SARS-CoV-2) compared to SARS-CoV-1 [Internet]. Infectious Diseases (except HIV/AIDS); 2020 Mar [cited 2020 Nov 25]. Available from: http://medrxiv.org/lookup/doi/10.1101/2020.03.09.20033217

8. Cai J, Sun W, Huang J, Gamber M, Wu J, He G. Indirect Virus Transmission in Cluster of COVID-19 Cases, Wenzhou, China, 2020. Emerg Infect Dis. 2020 Jun;26(6):1343–5.

9. Jia J, Ding J, Liu S, Liao G, Li J, Duan B, et al. Modeling the Control of COVID-19: Impact of Policy Interventions and Meteorological Factors. arXiv:200302985 [math, q-bio] [Internet]. 2020 Mar 5 [cited 2020 Nov 25]; Available from: http://arxiv.org/abs/2003.02985

10. Koo JR, Cook AR, Park M, Sun Y, Sun H, Lim JT, et al. Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. The Lancet Infectious Diseases. 2020 Jun;20(6):678–88.

11. Peng L, Yang W, Zhang D, Zhuge C, Hong L. Epidemic analysis of COVID-19 in China by dynamical modeling [Internet]. Epidemiology; 2020 Feb [cited 2020 Nov 25]. Available from: http://medrxiv.org/lookup/doi/10.1101/2020.02.16.20023465

12. Shen M, Peng Z, Xiao Y, Zhang L. Modeling the Epidemic Trend of the 2019 Novel Coronavirus Outbreak in China. The Innovation. 2020 Nov;1(3):100048.

13. Li D, Liu Z, Liu Q, Gao Z, Zhu J, Yang J, et al. Estimating the Efficacy of Quarantine and Traffic Blockage for the Epidemic Caused by 2019-nCoV (COVID-19):A Simulation Analysis [Internet]. Epidemiology; 2020 Feb [cited 2020 Nov 25]. Available from: http://medrxiv.org/lookup/doi/10.1101/2020.02.14.20022913

14. Peak CM, Kahn R, Grad YH, Childs LM, Li R, Lipsitch M, et al. Modeling the Comparative Impact of Individual Quarantine vs. Active Monitoring of Contacts for the Mitigation of COVID-19 [Internet]. Infectious Diseases (except HIV/AIDS); 2020 Mar [cited 2020 Nov 25]. Available from: http://medrxiv.org/lookup/doi/10.1101/2020.03.05.20031088

15. Ferguson N, Laydon D, Nedjati Gilani G, Imai N, Ainslie K, Baguelin M, et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand [Internet]. Imperial College London; 2020 Mar [cited 2020 Nov 25]. Available from: http://spiral.imperial.ac.uk/handle/10044/1/77482

16. Tang Z, Li X, Li H. Prediction of New Coronavirus Infection Based on a Modified SEIR Model. medRxiv [Internet]. 2020 Mar 6; Available from: https://www.medrxiv.org/content/10.1101/2020.03.03.20030858v1

17. Flaxman S, Mishra S, Gandy A, Unwin H, Coupland H, Mellan T, et al. Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries [Internet]. Imperial College London; 2020 Mar [cited 2020 Nov 25]. Available from: http://spiral.imperial.ac.uk/handle/10044/1/77731

18. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet. 2020 Jan 31;395:689–97.

19. Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science. 2020 May 1;368(6490):489–93.

20. Mizumoto K, Kagaya K, Zarebski A, Chowell G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. Eurosurveillance [Internet]. 2020 Mar 12 [cited 2020 Nov 25];25(10). Available from: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.10.2000180

21. Nishiura H, Kobayashi T, Miyama T, Suzuki A, Jung S-M, Hayashi K, et al. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). Int J Infect Dis. 2020;94:154–5.

22. Du Z, Xu X, Wu Y, Wang L, Cowling BJ, Meyers LA. Serial Interval of COVID-19 among Publicly Reported Confirmed Cases. Emerg Infect Dis. 2020 Jun;26(6):1341–3.

23. Ganyani T, Kremer C, Chen D, Torneri A, Faes C, Wallinga J, et al. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. Eurosurveillance [Internet]. 2020 Apr 30 [cited 2020 Nov 25];25(17). Available from: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.17.2000257

24. Chisholm RH, Campbell PT, Wu Y, Tong SYC, McVernon J, Geard N. Implications of asymptomatic carriers for infectious disease transmission and control. R Soc open sci. 2018 Feb;5(2):172341.

25. Barabási A-L, Pósfai M. Network science. Cambridge, United Kingdom: Cambridge University Press; 2016. 456 p.

26. WHO. Statement on the first meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV) [Internet]. 2020. Available from: https://www.who.int/news/item/23-01-2020-statement-on-the-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov)

27. Tang B, Wang X, Li Q, Bragazzi NL, Tang S, Xiao Y, et al. Estimation of the Transmission Risk of the 2019-nCoV and Its Implication for Public Health Interventions. JCM. 2020 Feb 7;9(2):462.

28. Liu T, Hu J, Xiao J, He G, Kang M, Rong Z, et al. Time-varying transmission dynamics of Novel Coronavirus Pneumonia in China [Internet]. Systems Biology; 2020 Jan [cited 2020 Nov 25]. Available from: http://biorxiv.org/lookup/doi/10.1101/2020.01.25.919787

29. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. N Engl J Med. 2020 Mar 26;382(13):1199–207.

30. Linton NM, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov AR, Jung S-M, et al. Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data. J Clin Med. 2020 Feb 17;9(2).

31. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. Annals of Internal Medicine. 2020 May 5;172(9):577–82.

32. Read JM, Bridgen JRE, Cummings DAT, Ho A, Jewell CP. Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions [Internet]. Infectious Diseases (except HIV/AIDS); 2020 Jan [cited 2020 Nov 25]. Available from: http://medrxiv.org/lookup/doi/10.1101/2020.01.23.20018549

33. Tian H, Liu Y, Li Y, Wu C-H, Chen B, Kraemer MUG, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science. 2020 May 8;368(6491):638–42.

34. Patrozou E, Mermel LA. Does Influenza Transmission Occur from Asymptomatic Infection or Prior to Symptom Onset? Public Health Rep. 2009 Mar;124(2):193–6.

35. Longini IM. Containing Pandemic Influenza with Antiviral Agents. American Journal of Epidemiology. 2004 Apr 1;159(7):623–33.

36. GLEAMviz [Internet]. Available from: http://www.gleamviz.org

752 37. Broeck WV den, Gioannini C, Gonçalves B, Quaggiotto M, Colizza V, Vespignani A. The
753 GLEaMviz computational tool, a publicly available software to explore realistic epidemic
754 spreading scenarios at the global scale. BMC Infect Dis. 2011 Dec;11(1):37.

755 38. Balcan D, Hu H, Goncalves B, Bajardi P, Poletto C, Ramasco JJ, et al. Seasonal transmission
756 potential and activity peaks of the new influenza A(H1N1): a Monte Carlo likelihood analysis
757 based on human mobility. BMC Med. 2009 Dec;7(1):45.

758 39. Balcan D, Gonçalves B, Hu H, Ramasco JJ, Colizza V, Vespignani A. Modeling the spatial
759 spread of infectious diseases: The GLobal Epidemic and Mobility computational model. Journal
760 of Computational Science. 2010 Aug;1(3):132–45.

761 40. Balcan D, Colizza V, Goncalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility
762 networks and the spatial spreading of infectious diseases. Proceedings of the National Academy
763 of Sciences. 2009 Dec 22;106(51):21484–9.

764 41. Simini F, González MC, Maritan A, Barabási A-L. A universal model for mobility and migration
765 patterns. Nature. 2012 Apr;484(7392):96–100.

766 42. Google. Community Mobility Reports [Internet]. Google; 2020. Available from:
767 https://www.google.com/covid19/mobility/

768 43. Waze. Waze Stats: Active Wazers in HU Budapest. [Internet]. Available from:
769 http://wazestats.com/active.php?city=7

770 44. COVID-19 pandemic in Hungary [Internet]. Available from: koronavirus.gov.hu

771 45. OAG. Coronavirus airline schedules data. [Internet]. Official Aviation Guide; 2020. Available
772 from: https://www.oag.com/coronavirus-airline-schedules-data

773 46. Ren Y, Ercsey-Ravasz M, Wang P, González MC, Toroczkai Z. Predicting commuter flows in
774 spatial networks using a radiation model based on temporal ranges. Nat Commun. 2014
775 Dec;5(1):5347.

776 47. Monto AS, DeJonge PM, Callear AP, Bazzi LA, Capriola SB, Malosh RE, et al. Coronavirus
777 Occurrence and Transmission Over 8 Years in the HIVE Cohort of Households in Michigan. The
778 Journal of Infectious Diseases. 2020 Jun 16;222(1):9–16.

779 48. KSH. H-UNCOVER – Reprezentatív felmérés a koronavírus elleni küzdelemben – eredmények
780 [Internet]. KSH; Available from:
781 http://www.ksh.hu/huncover_reprezentativ_felmeres_eredmenyek

782 49. Official COVID19 Dashboard public information [Internet]. Federal Ministry Republic of
783 Austria, Social Affairs, Health, Care and Consumer Protection; 2020. Available from:
784 https://info.gesundheitsministerium.at/?l=en

785 50. COVID-19: epidemiological update of March 4, 2020. [Internet]. Public Health France; 2020
786 Mar. Available from: https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-
787 19/coronavirus-chiffres-cles-et-evolution-de-la-covid-19-en-france-et-dans-le-monde#block-
788 266156

51. Coronavirus Disease 2019 (COVID-19) Daily Situation Report of the Robert Koch Institue. [Internet]. Robert Koch Institue; 2020 Mar. Available from: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/2020-03-04-en.pdf?__blob=publicationFile

52. Rosini. Cases tested people tested / tested cases [github] [Internet]. github; Available from: https://github.com/pcm- dpc/COVID-19/blob/master/dati-regioni/dpc-covid19-ita-regioni-20200304.csv

53. Report on confirmed COVID-19 cases in Spain. [Internet]. Carlos III. Health Institute; 2020 Mar. Available from: https://cnecovid.isciii.es/covid19/#ccaa

54. Iran COVID-19 death now 77 as emergency services chief infected. Channel New Asia [Internet]. 2020 Mar 4; Available from: https://www.channelnewsasia.com/news/world/covid19-coronavirus-death-toll-iran- 77-mar-3-12496192

55. Reustle S. Japan COVID-19 Coronavirus Tracker. [Internet]. Available from: https://covid19japan.com

56. MOH, Singapore. ONE MORE CASE DISCHARGED; TWO NEW CASES OF COVID-19 INFECTION CONFIRMED [Internet]. Singapore: Ministry of Health; 2020 Mar. Available from: https://www.moh.gov.sg/news-highlights/details/one-more-case-discharged-two-new-cases-of-covid-19-infection-confirmed

57. No. of COVID-19 Cases in S. Korea Rises to 5,621 [Internet]. KBS World Radio; 2020 Mar. Available from: http://world.kbs.co.kr/service/news_view.htm?lang=e&Seq_Code=151776

58. Pastor-Satorras R, Vespignani A. Epidemic Spreading in Scale-Free Networks. Phys Rev Lett. 2001 Apr 2;86(14):3200–3.

59. Meyers LA, Pourbohloul B, Newman MEJ, Skowronski DM, Brunham RC. Network theory and SARS: predicting outbreak diversity. Journal of Theoretical Biology. 2005 Jan;232(1):71–81.

60. Sapiezynski P, Stopczynski A, Lassen DD, Lehmann S. Interaction data from the Copenhagen Networks Study. Sci Data. 2019 Dec;6(1):315.

61. WHO. Considerations for quarantine of contacts of COVID-19 cases [Internet]. 2020 Aug. Available from: https://www.who.int/publications/i/item/considerations-for-quarantine-of-individuals-in-the-context-of-containment-for-coronavirus-disease-(covid-19)

62. Stopczynski A, Pentland A 'Sandy', Lehmann S. How Physical Proximity Shapes Complex Social Networks. Sci Rep. 2018 Dec;8(1):17722.

63. Berthold MR, Cebron N, Dill F, Gabriel TR, Kötter T, Meinl T, et al. KNIME: The Konstanz Information Miner. In: Preisach C, Burkhardt H, Schmidt-Thieme L, Decker R, editors. Data Analysis, Machine Learning and Applications. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008. p. 319–26.

64. Csárdi G, Nepusz T. Igraph - The network analysis package [Internet]. 2006. Available from: https://igraph.org

65. Albert R, Jeong H, Barabási A-L. Error and attack tolerance of complex networks. Nature. 2000 Jul 27;406(6794):378–82.

66. Feld SL. Why Your Friends Have More Friends Than You Do. American Journal of Sociology. 1991 May;96(6):1464–77.

67. Cohen R, Havlin S, ben-Avraham D. Efficient Immunization Strategies for Computer Networks and Populations. Phys Rev Lett. 2003 Dec 9;91(24):247901.

68. Knime workflow [Internet]. Available from: https://univet.hu/en/research/dfi/publications/scientific-publications/covid-19/

69. Adamecz-Völgyi A, Szabó-Morvai Á. Kik dolgoznak a frontvonalban? [Internet]. MTA Közgazdaság- és Regionális Tudományi Kutatóközpontja; 2020 Apr. Available from: https://www.mtakti.hu/koronavirus/kik-dolgoznak-a-frontvonalban/13090/

70. Miller A, Reandelar MJ, Fasciglione K, Roumenova V, Li Y, Otazu GH. Correlation between universal BCG vaccination policy and reduced mortality for COVID-19. medRxiv. 2020 Sep 14;

71. Berg MK, Yu Q, Salvador CE, Melani I, Kitayama S. Mandated Bacillus Calmette-Guérin (BCG) vaccination predicts flattened curves for the spread of COVID-19 [Internet]. Public and Global Health; 2020 Apr [cited 2020 Nov 25]. Available from: http://medrxiv.org/lookup/doi/10.1101/2020.04.05.20054163

72. HUNgarian COronaVirus Disease-19 Epidemiological Research [Internet]. Semmelweis University Heart and Vascular Center; 2020 Jun. Available from: https://clinicaltrials.gov/ct2/show/study/NCT04370067

73. Oliver N, Lepri B, Sterly H, Lambiotte R, Deletaille S, De Nadai M, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. Sci Adv. 2020 Jun;6(23):eabc0764.

1 Figure captions

Figure 1.

Structure of the modified SEIR model with compartment initials (Susceptible, Exposed (Latent), Infected, Recovered), transitions from one compartment to another (full line), parameters and routes of infection (dashed line). Susceptible individuals can get the infection either from individuals being in the Infected compartment with the rate of β, or from Exposed compartment (latent individuals) with the rate of β_L.

Figure 2.

Distribution of Infected individuals in *Scenarios 1* to *3.*

Figure 3.

Distribution of Exposed (latent) individuals in *Scenarios 1* to *3.*

Figure 4.

Comparison of the distribution of Infected individuals in case of modified SEIR model and basic SEIR model in case of *Scenario 1*.

Figure 5.

Results of sensitivity analysis of the modified SEIR model: changes in daily maximum value of individuals in the infected compartment (A) and changes in the number of the days related to the former endpoint (B)

Figure 6.

Epidemic thresholds (λ'_c) of different degree exponents (γ) and target degrees (k_t) calculated to increase the epidemic threshold above the spreading rate of COVID-19 with targeted sampling. The spreading rate (λ) of the virus is set to 0.067 (red line).