*Article*

# Risk assessment of novel coronavirus 2019-nCoV outbreaks outside China

**Péter Boldog[1], Tamás Tekeli[1], Zsolt Vizi[1], Attila Dénes[1]*, Ferenc A. Bartha[1] and Gergely Röst[1]**

[1] Bolyai Institute, University of Szeged, Hungary
* Correspondence: denesa@math.u-szeged.hu (A.D.)

Received: date; Accepted: date; Published: date

**Abstract:** We developed a computational tool to assess the risk of novel coronavirus outbreaks outside China. We estimate the dependence of the risk of a major outbreak in a country from imported cases on key parameters such as: (i) the evolution of the cumulative number of cases in mainland China outside the closed areas; (ii) the connectivity of the destination country with China, including baseline travel frequencies, the effect of travel restrictions, and the efficacy of entry screening at destination; (iii) the efficacy of control measures in the destination country (expressed by the local reproduction number $R_{\text{loc}}$). We found that in countries with low connectivity to China but with relatively high $R_{\text{loc}}$, the most beneficial control measure to reduce the risk of outbreaks is a further reduction in their importation number either by entry screening or travel restrictions. Countries with high connectivity but low $R_{\text{loc}}$ benefit the most from policies that further reduce $R_{\text{loc}}$. Countries in the middle should consider a combination of such policies. Risk assessments were illustrated for selected groups of countries from America, Asia and Europe, and we investigated how their risks depend on those parameters, and how the risk is increasing in time as the number of cases in China is growing.

**Keywords:** novel coronavirus, transmission, risk assessment, interventions, travel, outbreak, nCoV-2019, compartmental model, branching process

## 1. Introduction

A cluster of pneumonia cases in Wuhan, China, was reported to World Health Organization (WHO) on 31st December 2019. The cause of the pneumonia cases was identified as a novel betacoronavirus, the 2019 novel coronavirus (2019-nCoV). The first patient showing symptoms was recorded by Chinese authorities on 8th December 2019 [1]. On 9th January 2020, WHO confirmed that a novel coronavirus had been isolated from one of the hospitalised persons [2], and the first death case was reported on the same day. The first case outside China was witnessed on 13th January in Thailand [3], and in the following days, several other countries also reported 2019-nCoV cases [4]. The first confirmed cases in China, but outside Hubei province, were reported on 19th January. [4]. As of 1st February, there were 14628 confirmed cases worldwide (out of which 14451 happened in China) with 305 total deaths [5].

Since no specific antiviral agent is available for treatment of this infection, and there is no vaccine [6], the control measures, introduced both in China and other countries, aimed to prevent the transmission. A metropolitan-wide quarantine of Wuhan and nearby cities was introduced on 23–24 January [7]. Several airports and train stations have started temperature screening measures to identify people with fever [8]. All public transportation was suspended in Wuhan from 10 a.m., 23rd January, including all outbound train and flights as well as all bus, metro and ferry lines, additionally, all outbound trains and flights were halted [9]. Construction of a specialist emergency hospital was started in Wuhan [10], and nearly 6000 medical workers were sent to Wuhan from across China [11]. Beijing also announced the suspension of all inter-provincial bus and train services, several touristic attractions, including the Forbidden City and the Shanghai Disneyland were closed [9]. Other countries also introduced control measures, including screening passengers arriving from China and closing

their borders [12]. Several airlines, including British Airways and Lufthansa cancelled all flights to and from mainland China [9].

The potential dangers of 2019-nCoV have prompted a number of studies on its epidemiological characteristics. The 2018 travel data from International Air Transport Association (IATA) was used to identify the countries, and their Infectious Disease Vulnerability Index (IDVI) [13], that receive substantial travel inflow from Wuhan Tianhe International Airport [14]. The IDVI has a range of 0–1 with higher score implying lower vulnerability. The top destinations Bangkok, Hong Kong, Tokyo and Taipei all have an IDVI above 0.65.

It is essential to estimate the number of infections (including those that have not been diagnosed), to be able to analyze the spread of the disease. For 17$^{th}$ January 2020, preliminary estimates were given as 1250 (range: 350–3000) [15] for the total number of infections up to that date. Later, Imai *et al.* [16] estimated the total number of infections in China and warned that their number is likely to substantially exceed that of the officially confirmed cases (see also [17]). They report an estimate of 4000 infections (range: 1000–9700) by 18$^{th}$ January 2020. Analyzing exported cases, Nishiura *et al.* obtained comparable numbers that is 5502 (range: 3027–9057) infections by 24$^{th}$ January 2020 [18]. Estimates from exportations and an individual based mobility model were obtained by Chinazzi *et al.* [15].

To better assess the epidemic risk of 2019-nCoV, among the key parameters to be approximated are the basic reproduction number $R_0$ and the incubation period. We summarize previous efforts to this aim in Table 1, and present a short summary below.

The majority of the estimates for $R_0$ range between 2–3. Obtaining these was done by modeling epidemic trajectories and comparing them to the results of [16] as a baseline [19,20], using a negative binomial distribution to generate secondary infections. Liu *et al.* utilized the Exponential Growth and Maximum Likelihood Estimation methods and found that the 2019-nCoV may have a higher pandemic risk than SARS-CoV in 2003 and [21].

Read *et al.* based their estimates on data from Wuhan exclusively (available up to 22$^{nd}$ January 2020) and a deterministic SEIR model [22]. The choice of this date is motivated by the actions of authorities that is the substantial travel limitations the next day. Li *et al.* used solely the patient data with illness onset between 10$^{th}$ December 2019 and 4$^{th}$ January 2020 [23]. The Centre for the Mathematical Modelling of Infectious Diseases at the London School of Hygiene & Tropical Medicine have analyzed 2019-nCoV using SEIR and multiple data series [24]. Shen *et al.* used a SEIJR model and Markov Chain Monte Carlo (MCMC) simulations [25] similar to [26]. An alternative approach was presented by Majumder and Mandl [27] as they obtained their estimate based on the cumulative epidemic curve and the Incidence Decay and Exponential Adjustment (IDEA) model [28].

The incubation period was estimated to be in between 4.6–5.8 days by various studies. First calculations used data up to 23$^{rd}$ January [21]. Backer *et al.* used newly available patient data with known travel history [29]. Linton *et al.* gave estimates for with and without Wuhan residents using their statistical model with Weibull distribution [30]. The Johns Hopkins University Infectious Disease Dynamics Group has been collecting substantial data on exposure and symptom onset for 2019-nCoV cases. They recommend using their LogNormal estimate [31], which gives a $\sim$ 5.1 days incubation period.

**Table 1.** Published estimates of the key epidemiological parameters of 2019-nCoV
uncertainty range is given where provided

| $R_0$ | Incubation period | Method of estimation | Reference |
|---|---|---|---|
| 2.6 (1.5–3.5) | - | Epidemic Simulations | [19] |
| 2.2 (1.4–3.8) | - | Stochastic Simulations | [20] |
| 2.9 (2.3–3.6) | 4.8 days | Exp. Growth, Max. Likelihood Est. | [21] |
| 2.56 (2.49–2.63) | - | Exp. Growth, Max. Likelihood Est. | [17] |
| 3.11 (2.3–4.1) | - | SEIR | [22] |
| 2.5 (2.0–3.1) | - | Incidence Decay and Exponential Adjustment model | [27] |
| 2.2 (1.4–3.9) | 5.2 days (4.1–7.0) | Renewal Equations | [23] |
| * (1.4–4.0) | - | SEIR | [24] |
| 4.71 (4.5–4.9) | 5.0 days (4.9–5.1) | Dec. 2019, SEIJR, MCMC | [25] |
| 2.08 (1.9–2.2) | | Jan. 2020, SEIJR, MCMC | [25] |
| 2.68 (2.4–2.9) | - | SEIR, MCMC | [26] |
| - | 5.8 days (4.6–7.9) | Weibull | [29] |
| - | 4.6 days (3.3–5.8) | Weibull incl. Wuhan | [30] |
| - | 5.0 days (4.0–5.8) | Weibull excl. Wuhan | [30] |
| - | 5.1 days (4.4–6.1) | LogNormal | [31] |

In this study we combine case estimates, epidemiological characteristics of the disease, international mobility patterns, control efforts, and secondary case distributions to assess the risk of major outbreaks from imported cases outside China.

## 2. Materials and Methods

### *2.1. Model ingredients*

Our method has three main components:

(i) We estimate the cumulative number of cases in China outside Hubei province after 23rd January, using a time-dependent compartmental model of the transmission dynamics.
(ii) We use that number as an input to the global transportation network to generate probability distributions of the number of infected travellers arriving at destinations outside China.
(iii) In a destination country, we use a Galton–Watson branching process to model the initial spread of the virus. We calculate the extinction probability of each branch initiated by a single important case, obtaining the probability of a major outbreak as the probability that at least one branch will not extinct.

### *2.2. Epidemic size in China outside the closed areas of Hubei*

The starting point of our transmission model is 23rd January, when major cities in Hubei province were closed [7]. From this point forward, we run a time dependent $SE_nI_mR$ model in China outside Hubei, which was calibrated to be consistent with the estimated case numbers outside Hubei until 31st January. We impose time dependence in the transmission parameter due to the control measures progressively implemented by Chinese authorities on and after 23rd January. With our baseline $R_0 = 2.6$, disease control is achieved when more than 61.5% of potential transmissions are prevented. We introduce a key parameter $t_*$ to denote the future time when control measures reach their full potential. For this study we assume it to be in the range of 20–50 days after 23rd January. Using our transmission model, we calculate the total cumulative number of cases (epidemic final size) outside Hubei, for each $t_*$ in the given range. This also gives an upper bound for the increasing cumulative number of cases $C = C(t)$.

### *2.3. Connectivity and case exportation*

The output $C$ of the transmission model is used as the pool of potential travellers to abroad, and fed into the online platform EpiRisk [32]. This way, similarly as in [15], we evaluated the probability

that a single infected individual is traveling from the index areas (in our case Chinese provinces other than Hubei) to a specific destination. This baseline probability can be multiplied by other factors such as the reduction in travel volume between the index and destination area, and the efficacy of entry screening at the destination country. Hence we have a compound parameter, denoted by $\theta$, that expresses the probability that a case in China outside Hubei will be eventually mixed into the population of the destination country. For example, the probability that a case from China ends up in Japan is $2.42 \times 10^{-4}$, under normal circumstances, during the January–February period [32]. Assuming a 20% reduction in travel volume between China and Japan, and a 40% efficacy on entry screening [33], the connectivity parameter that we use in our model becomes $\theta = 1.16 \times 10^{-4}$. Assuming independence, this $\theta$, together with $C$ generates a binomial distribution of importations that enter the population of a given country.

### *2.4. Probability of a major outbreak in a country by imported cases*

Each imported case that passed the entry screening and mixed into the local population can potentially start an outbreak, what we model by a Galton–Watson branching process with negative binomial offspring distribution with dispersion parameter $k = 0.64$ [19,20] and expectation $R_{\text{loc}}$, where $R_{\text{loc}}$ is the local reproduction number of the infection in a given country. Each branch has extinction probability $z$, which is the unique solution of the equation $z = g(z)$ on the interval $(0, 1)$ , where $g$ is the generating function of the offspring distribution. The process dies out if all the branches die out, thus we estimate the risk of a major local outbreak from importation as $1 - z^i$, where $i$ cases were imported.

### *2.5. Dependence of the risk of major outbreaks on key parameters*

The number of imported cases $i$ is given by a random variable $X$, where $X \sim \text{Binom}(C, \theta)$. The outbreak risk in a country $x$ is then estimated as $\text{Risk}_x = E[1 - z^X]$, where $E$ is the expectation of the outbreak probabilities, thus we consider a probability distribution of branching processes. This way $\text{Risk}_x = \text{Risk}(C, \theta, R_{\text{loc}})$, which means that the risk depends on the efficacy of Chinese control measures that influence the cumulative case number $C$, the connectivity between the index and destination areas $\theta$, and the local reproduction number $R_{\text{loc}}$. The main question we aim to get insight into is how this risk depends on these three determining factors.

The technical details of the modelling and calculations can be found in Appendices A, B, and C.

## 3. Results

### *3.1. Epidemic size in China*

After calibration of the $SE_2I_3R$ model, we numerically calculated the final epidemic size (total cumulative number of cases) in China outside Hubei, using three different basic reproduction numbers and different control functions. The control functions were parametrised by $t_*$, which is the time after 23$^{\text{rd}}$ January, when the control reaches its maximal value $u_{\max}$. Smaller $t_*$ corresponds to more rapid implementation of the control measures. In Figure 1, we plotted these cumulative numbers versus $t_*$, and we can observe that the epidemic final size is rather sensitive to the speed of implementation of the control measures. These curves also give upper bounds for the number of cumulative cases at any given time, assuming that the control efforts will be successful.

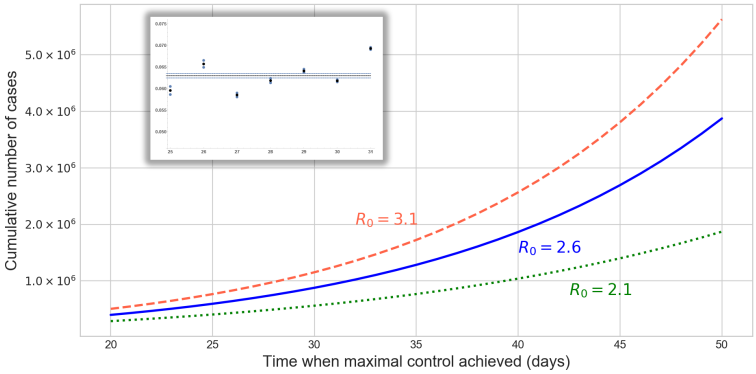


**Figure 1.** Final epidemic sizes in China, outside Hubei, with $R_0 = 2.1, 2.6, 3.1$, as a function of the time when the control function $u(t)$ reaches its maximum (in days after 24$^{\text{th}}$ January). Rapid implementation of the control generates much smaller case numbers. The inset shows the estimations of the ascertainment rate between 24$^{\text{th}}$ and 31$^{\text{st}}$, with average 0.063, based on the ratio of confirmed cases and the maximum likelihood estimates of the case numbers from exportations.

### 3.2. *Risk of major outbreaks*

We generated a number of plots to depict Risk$(C, \theta, R_{\text{loc}})$ for selected groups of countries from America, Europe and Asia.

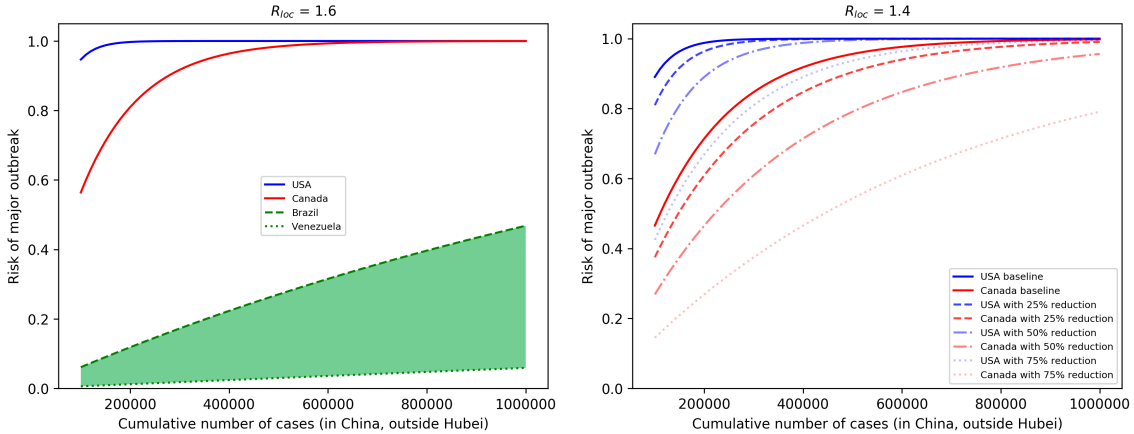


**Figure 2.** In the left: Risk of major outbreaks as a function of cumulative number of cases in selected countries, assuming $R_{\text{loc}} = 1.6$ and baseline connectivity to China. Other countries in South America, and Mexico are inside the green shaded area. In the right: the effect of the reduction of imported case numbers (either by travel restriction or entry screening) in the USA and Canada, assuming $R_{\text{loc}} = 1.4$.

In the left of Figure 2, we can see the risk of American countries as a function of cumulative number of cases $C$, assuming each country has $R_{\text{loc}} = 1.6$ and their connectivity is their baseline $\theta$. When $C$ exceeds 600 000, with this local reproduction number and without any restriction in importation, outbreaks in the USA and Canada are very likely, while countries in South America (including Mexico), which are all in the green shaded region, still have moderate risk. To illustrate the impact of control measures for USA and Canada, we reduced $R_{\text{loc}}$ to 1.4, and plotted the risk for different levels of reduction in connectivity to China, either due to travel restrictions or entry screening, see Figure 2 in the right. As the number of cases in China approaches one million, such reductions have limited effect on the risk of outbreak. Figure 1 provides us with scenarios when $C$ remains below certain values.

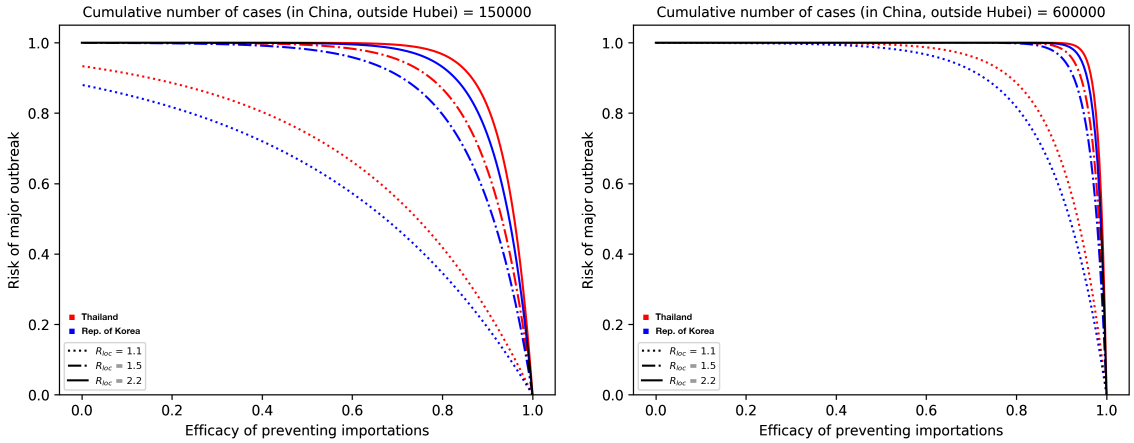


**Figure 3.** Outbreak risks for highly connected countries in Asia. Thailand and Rep. Korea are plotted, the curves for Japan and Taiwan are in between them. In the left, we plot the risk vs. the efficacy of prevented importations when the cumulative number of cases reaches 150000. In the right, $C = 600000$. Black parts of the curves represent situations when the four countries are indistinguishable.

We considered the group of countries from Asia which are the most connected to China: Thailand, Japan, Taiwan and Rep. Korea. They have similar baseline connectivity $\theta$, and we focus on how travel restrictions and entry screenings can potentially reduce their risks, assuming different values of $R_{\text{loc}}$ in the case $C = 150000$ (in the left of Figure 3) and $C = 600000$ (in the right of Figure 3). For illustration purposes we plotted Thailand (red) and Rep. Korea (blue), but Taiwan and Japan are always between those two curves. We can see that, for example, in the right of Figure 3 for $C = 600000$ , unless $R_{\text{loc}}$ is very small, considerable reduction of the outbreak risk can be achieved only by extreme measures that prevent most importations.

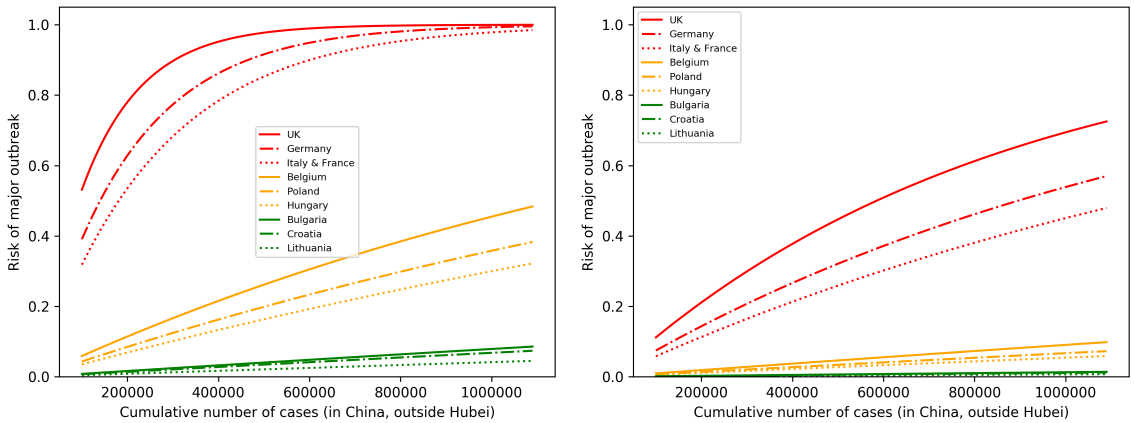


**Figure 4.** Selected European countries with high, medium and low connectivity to China. In the left, the outbreak risk is plotted assuming their baseline connectivity $\theta$, and $R_{\text{loc}} = 1.4$ for each country, as the cumulative number of cases is increasing. A significant reduction in the risks can be observed in the right, where we reduced $R_{\text{loc}}$ to 1.1 and assumed a 50% reduction in importations.

In Figure 4, we assumed that European countries have very similar $R_{\text{loc}}$ and looked at their risks as a function of the number of cases. For illustration purposes, we selected countries which have relatively high (UK, Germany, France, Italy), medium (Belgium, Poland, Hungary) and low (Bulgaria, Croatia, Lithuania) connectivity to China. In the left, we assumed $R_{\text{loc}} = 1.4$ and baseline $\theta$, and with these parameters, outbreaks will likely occur in high risk countries as the case number approaches one million. By reducing $R_{\text{loc}}$ to 1.1 and the by reducing $\theta$ to the half of its baseline (meaning that we

assume that there is a 50% reduction in importations due to decreased travel and entry screenings), then the risk is significantly reduced even with one million cases.

### *3.3. Profile of countries benefiting the most from interventions*

We also plotted the risk on a two-parameter map, as a function of $\theta$ and $R_{\text{loc}}$. Observing the gradients of the risk map, we can conclude that countries with low connectivity but high $R_{\text{loc}}$ should focus on further reducing importations by entry screening and travel restrictions, while countries with high connectivity but smaller $R_{\text{loc}}$ better focus on control measures that potentially further reduce $R_{\text{loc}}$. Countries in the middle benefit most from the combination of those two types of measures.

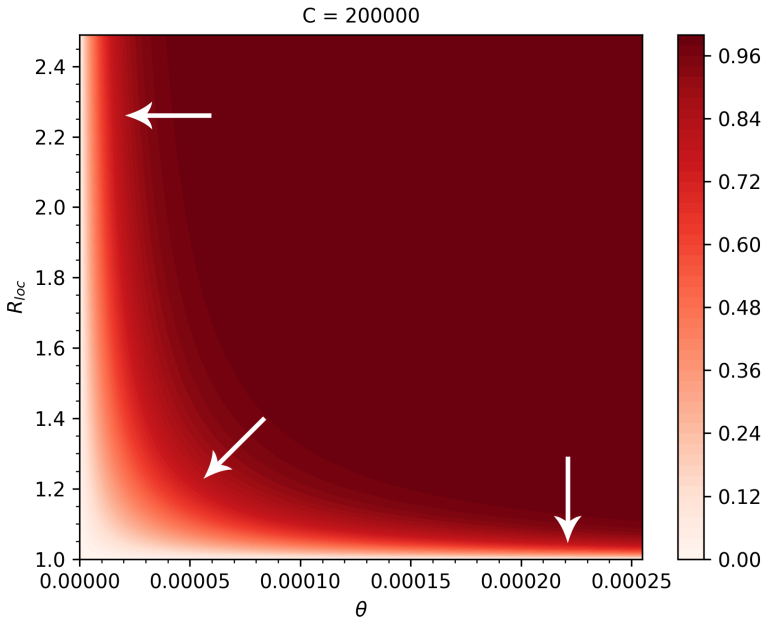


**Figure 5.** Heatmap of the outbreak risks as a function of $\theta$ and $R_{\text{loc}}$, when $C = 200000$. The arrows show the directions corresponding to the largest reduction in the risk.

## 4. Discussion

By combining three different modelling approaches, we created a tool to assess the risk of 2019-nCoV outbreaks in countries outside China. This risk depends on three key parameters: the cumulative number of cases in areas of China which are not closed, the connectivity between China and the destination country, and the local transmission potential of the virus. Quantifications of the outbreak risks and their dependence on the key parameters were illustrated for selected groups of countries from America, Asia and Europe, representing a variety of country profiles.

There are several limitations of our model, as each ingredient use assumptions, which are detailed in the Appendices. There is a great uncertainty in the epidemiological parameters as well. It is difficult to predict the epidemic trajectory in China, as the effects of the control measures are not clear yet. There were recent disruptions in international travel, suggesting that the EpiRisk parameters will not be accurate in the future. Nevertheless, when we will have new information in the future about the case numbers in China, travel frequencies, efficacy of entry screenings and local control measures, our method will still be useful to assess outbreak risks.

We found that in countries with low connectivity to China but with relatively high $R_{\text{loc}}$, the most beneficial control measure to reduce the risk of outbreaks is a further reduction in their importation number either by entry screening or travel restrictions (see Figure 5). Countries with high connectivity

but low $R_{\text{loc}}$ benefit the most from policies that further reduce $R_{\text{loc}}$. Countries in the middle should consider a combination of such policies.

Cumulative cases and connectivity can be estimated, in general. However, to make a good assessment of the outbreak risk, it is very important to estimate $R_{\text{loc}}$ in each country. In the absence of available transmission data, one may rely on the experiences from previous outbreaks. Knowing $R_{loc}$ is needed not only to have a better quantitative risk estimation, but also for guidance to what type of control measures may reduce the outbreak risk the most.

**Author Contributions:** Conceptualization and methodology, G.R.; codes and computations A.D., T.T., F.B., P.B., G.R., Zs.V.; data collection and analysis, A.D., F.B., T.T.; writing and editing, A.D., F.B., G.R.; visualization, P.B., T.T., Zs.V.

**Funding:** G.R. was supported by EFOP-3.6.1-16-2016-00008. F.B. was supported by NKFIH KKP 129877. T.T. was supported by NKFI FK 124016, A.D. was supported by NKFIH PD 128363 and by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. P.B. was supported by 20391-3/2018/FEKUSTRAT.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A Transmission dynamics

The governing system of the transmission dynamics model is

$$S' = -(1-u)\beta S \sum_{i=1}^{3} I_k/N, \quad E_1' = (1-u)\beta S \sum_{i=1}^{3} I_k/N - 2\alpha E_1, \quad E_2' = 2\alpha E_1 - 2\alpha E_2,$$

$$I_1' = 2\alpha E_2 - 3\gamma I_1 - \mu I_1, \quad I_2' = 3\gamma I_1 - 3\gamma I_2 - \mu I_2, \quad I_3' = 3\gamma I_2 - 3\gamma I_3 - \mu I_3, \quad R' = 3\gamma I_3.$$

This is an extension of a standard SEIR model assuming gamma-distributed incubation and infectious periods, with the Erlang-parameters $n = 2, m = 3$ (following the SARS-study [34]). Given that disease fatalities do not have significant effect on the total population, we ignored them in the transmission model to ease the calculations (i.e. $\mu = 0$ was used). In this model, the basic reproduction number is $R_0 = \beta/\gamma$ , the incubation period is $\alpha^{-1}$ and the infectious period is $\gamma^{-1}$. The model is used to describe the disease dynamics in China outside Hubei province after 23$^{\text{rd}}$ January. We assume that at time $t$ after 23$^{\text{rd}}$ January, an increasing control function $u(t)$ represents the fraction of the transmissions that are prevented, thus the effective reproduction number becomes $R(t) = (1-u(t))R_0 S(t)/N$.

Based on the previous estimates from the literature (see Table 1), we chose an incubation period $\alpha^{-1} = 5.1$ days [31], basic reproduction number $R_0 = 2.6$ (2.1—3.1) with the corresponding infectious period $\gamma^{-1} = 3.3$ (1.7-5.6) days [19]. To predict the final number of cases outside Hubei, we assume a gradually increasing control $u$ from zero until a saturation point, and define $t_*$ the time when the eventual control $u_{\max}$ is achieved. The sooner this happens, the more successful the control is. Using the control term $u(t) = \min\{u_{\max} t/t_*, u_{\max}\}$, disease control is reached at $t = t_*(1 - 1/R_0)/u_{\max}$. For the calculations we choose $u_{\max} = 0.8$, noting that such a drop in transmission has been observed for SARS, where the reproduction number was largely reduced by subsequent interventions [35]. With our baseline $R_0 = 2.6$, disease control $R(t) < 1$ is achieved when $u(t) > 0.615$, meaning that more than 61.5% of potential transmissions are prevented, which occurs at time $t = 0.77 t_*$.

Since the first case outside Hubei was reported on 19$^{\text{th}}$ January [5], for the initialization of the model we could assume that number of infected individuals on 23$^{\text{rd}}$ January outside Hubei was equal to the number of cumulative cases outside Hubei up to that day. To calibrate the model, we estimated the number of cases from 24$^{\text{th}}$ January till 31$^{\text{st}}$ January outside Hubei based on case exportations, using the methodology of [15] , assuming that exportations after 24$^{\text{th}}$ January were only from outside Hubei. Based on the maximal likelihood of case numbers that produce the observed number of exportations using EpiRisk [32], we estimate that the reported confirmed cases represent only 6.3% of the total cases for the regions outside Hubei (other estimates for ascertainment rate were: 5.1% [22], 10% in [36], and 9.2% (95% confidence interval: 5.0, 20.0) [37]), see the inset in Figure 1. The initial values for the exposed compartments in the SEIR model were selected such that the model output was consistent

with the estimated case numbers outside Hubei between 24$^{th}$ January and 31$^{st}$ January. Solving the compartmental model, we obtained final epidemic sizes for various reproduction numbers and control efforts (see Figure 1), providing upper bounds for the cumulative number of cases $C$ outside Hubei.

## Appendix B Calculating the risk of outbreaks by importation

We create a probabilistic model to estimate the risk of a major outbreak in a destination country as a function of the cumulative number of cases $C$ in China outside the closed areas, the local reproduction number $R_{\text{loc}}$ in the destination country, and the connectivity $\theta$ between China and the destination country. We summarize these in Table B1.

**Table B1.** Parameters for calculating the risk of major outbreaks

| Parameter | Interpretation | Depends on ... | Typical Range |
|---|---|---|---|
| $C$ | Cumulative case number<br>in China, outside the closed areas | properties of nCoV-2019,<br>efficacy of Chinese control | $[100K, 6000K]$ |
| $R_{\text{loc}}$ | Local reproduction number<br>in destination country | destination country | $[1, 2.6]$ |
| $\theta$ | Probability of a importation<br>chance that a case from the origin travelling to and<br>mixing into the local population of the destination country | China and<br>destination country | $[0, 0.00025]$ |

We assume that the number of the imported cases entering the local population of the destination country follows a binomial distribution, i.e. the probability $p_i$ corresponding to $i$ imported cases in the destination country with connectivity $\theta$ to China is given by

$$p_i = \binom{C}{i} \theta^i (1-\theta)^{C-i}.$$

We calculate the extinction probability $z$ of a branching process initiated by a new infection in the destination country. As in [19,20], we assume the number of secondary infections to follow a negative binomial distribution with generator function

$$g(z) = \left( \frac{q}{1-(1-q)z} \right)^k,$$

with dispersion parameter $k$ and mean $\mu = R_{\text{loc}}$. Then, the probability parameter $q$ of the distribution is obtained as $q = \frac{k}{k+R_{\text{loc}}}$. The extinction probability of a branch is the solution of the fixed point equation $z = g(z)$.

Assuming that the destination country has $i$ imported cases from China that are mixed into the local population, we estimate the probability of a major outbreak as the probability that not all the branches started by those $i$ individuals die out, which is $1 - z^i$. Thus, the expectation of the risk of a major outbreak in country $x$ can be calculated as

$$\text{Risk}_x = \sum_{i=0}^{C} p_i (1 - z^i) = 1 - (\theta z + 1 - \theta)^C,$$

where we used the binomial theorem to simplify the sum. Having the input values of the parameters $C, \theta, R_{\text{loc}}$, with this model we can numerically calculate the risk.

## Appendix C

The codes for the computations were implemented in Mathematica and in Python, and they are available, including the used data, at [38].

## References


1. Huang, C. *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* **2020**. https://doi.org/10.1016/S0140-6736(20)30183-5.
2. WHO. Statement Regarding Cluster of Pneumonia Cases in Wuhan, China. *World Health Organization* **2020.01.09**. https://www.who.int/china/news/detail/09-01-2020-who-statement-regarding-cluster-of-pneumonia-cases-in-wuhan-china.
3. WHO. Novel Coronavirus – Thailand (ex-China). *World Health Organization* **2020.01.14**. https://www.who.int/csr/don/14-january-2020-novel-coronavirus-thailand-ex-china/en.
4. WHO. Novel Coronavirus (2019-nCoV) Situation Report – 1. *World Health Organization* **2020.01.21**. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf.
5. JHU IDD team. 2019-nCoV Global Cases by Center for Systems Science and Engineering. *JHU IDD* **2020**. https://docs.google.com/spreadsheets/d/1wQVypefm946ch4XDp37uZ-wartW4V7ILdg-qYiDXUHM/edit?usp=sharing.
6. CDC. 2019 Novel Coronavirus. Prevention & Treatment. *Centers for Disease Control and Prevention* **2020**. https://www.cdc.gov/coronavirus/2019-ncov/about/prevention-treatment.html.
7. NPR. Chinese Authorities Begin Quarantine Of Wuhan City As Coronavirus Cases Multiply. *NPR* **2020.01.23**. https://www.npr.org/2020/01/23/798789671/chinese-authorities-begin-quarantine-of-wuhan-city-as-coronavirus-cases-multiply.
8. Cheng, W.C.C.; Wong, S.-C.; To, K.K.W.; Ho, P.L.; Yuen, K.-Y. Preparedness and proactive infection control measures against the emerging Wuhan coronavirus pneumonia in China. *J. Hosp. Infect.* **2020**, to appear. https://doi.org/10.1016/j.jhin.2020.01.010.
9. Arnot, M.; Mzezewa, T. The Coronavirus: What Travelers Need to Know. *The New York Times* **2020.01.26**. https://www.nytimes.com/2020/01/26/travel/Coronavirus-travel.html.
10. National Health Commission of the People's Republic of China. Work begins on mobile hospital in Wuhan. *NHC* **2020.01.29**. http://en.nhc.gov.cn/2020-01/29/c_76034.htm.
11. National Health Commission of the People's Republic of China. Medics flood to Hubei to fight disease. *NHC* **2020.01.29**. http://en.nhc.gov.cn/2020-01/29/c_76031.htm.
12. Parry, J. Pneumonia in China: lack of information raises concerns among Hong Kong health workers. *BMJ* **2020**;368:m56. https://doi.org/10.1136/bmj.m56.
13. Moore, M.; Gelfeld, B.; Okunogbe, A. T.; Christopher, Paul. Identifying Future Disease Hot Spots: Infectious Disease Vulnerability Index. RAND Corporation: Santa Monica, CA, 2016. https://www.rand.org/pubs/research_reports/RR1605.html. Also available in print form.
14. Bogoch, I.I. *et al.* Pneumonia of Unknown Etiology in Wuhan, China: Potential for International Spread Via Commercial Air Travel. *J. Travel Med.* **2020**, accepted. https://doi.org/10.1093/jtm/taaa008.
15. Chinazzi, M. *et al.* Preliminary assessment of the International Spreading Risk Associated with the 2019 novel Coronavirus (2019-nCoV) outbreak in Wuhan City. *Laboratory for the Modeling of Biological and Socio–technical Systems* **2020.01.17**. https://www.mobs-lab.org/uploads/6/7/8/7/6787877/wuhan_novel_coronavirus__6_.pdf.
16. Imai, N. *et al.* Report 2: Estimating the potential total number of novel Coronavirus cases in Wuhan City, China. *Imperial College London* **2020.01.22**. https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/2019-nCoV-outbreak-report-22-01-2020.pdf.
17. Zhao, S. *et al.* Estimating the Unreported Number of Novel Coronavirus (2019-nCoV) Cases in China in the First Half of January 2020: A Data-Driven Modelling Analysis of the Early Outbreak. *J. Clin. Med.* **2020**, *9(2)*, 388. https://doi.org/10.3390/jcm9020388.
18. Nishiura, H. *et al.* The Extent of Transmission of Novel Coronavirus in Wuhan, China, 2020. *J. Clin. Med.* **2020**, 9(2), 330. https://doi.org/10.3390/jcm9020330.
19. Imai, N. *et al.* Report 3: Transmissibility of 2019-nCoV. *Imperial College London* **2020.01.24**. https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-2019-nCoV-transmissibility.pdf.
20. Riou, J.; Althaus, C.L. Pattern of early human-to-human transmission of Wuhan 2019-nCoV. *bioRχiv* **2020**. https://doi.org/10.1101/2020.01.23.917351.

21. Liu, T. *et al.* Transmission dynamics of 2019 novel coronavirus (2019-nCoV). *bioRχiv* **2020**. https://doi.org/10.1101/2020.01.25.919787.
22. Read, J.M.; Bridgen, J.R.E.; Cummings, D.A.T.; Ho, A.; Jewell, C.P. Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. *medRχiv* **2020**. https://doi.org/10.1101/2020.01.23.20018549.
23. Li, Q. *et al.* Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *N. Engl. J. Med.* **2020**. https://doi.org/10.1056/NEJMoa2001316.
24. Kucharski, A. *et al.* Analysis of early transmission dynamics of nCoV in Wuhan. *CMMID* **2020**. https://cmmid.github.io/ncov/wuhan_early_dynamics.
25. Shen, M.; Peng, Z.; Xiao, Y.; Zhang, L. Modelling the epidemic trend of the 2019 novel coronavirus outbreak in China. *bioRχiv* **2020**. https://www.biorxiv.org/content/10.1101/2020.01.23.916726v1.
26. Leung, K.; Wu, J.T.; Leung, G.M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet* **2020**, to appear. https://doi.org/10.1016/S0140-6736(20)30260-9.
27. Majumder M.; Mandl K.D. Early Transmissibility Assessment of a Novel Coronavirus in Wuhan, China. *SSRN* **2020**. https://dx.doi.org/10.2139/ssrn.3524675.
28. Fisman, D.N.; Hauck, T.S.; Tuite, A.R.; Greer, A.L. An IDEA for Short Term Outbreak Projection: Nearcasting Using the Basic Reproduction Number. *PLOS One* **2013**, *8(12)*. https://doi.org/10.1371/journal.pone.0083622.
29. Backer, J.A.; Klinkenberg, D.; Wallinga, J. The incubation period of 2019-nCoV infections among travellers from Wuhan, China. *medRχiv* **2020**. https://doi.org/10.1101/2020.01.27.20018986.
30. Linton, N.M. *et al.* Epidemiological characteristics of novel coronavirus infection: A statistical analysis of publicly available case data. *medRχiv* **2020**. https://doi.org/10.1101/2020.01.26.20018754.
31. Zheng, Q. *et al.* Real-time estimation of the novel coronavirus incubation time. *JHU IDD* **2020**. https://github.com/HopkinsIDD/ncov_incubation.
32. EpiRisk. Available from: http://epirisk.net.
33. Quilty, B.; Clifford, S.; CMMID nCoV working group; Flasche, S.; Eggo, R.M. Effectiveness of airport screening at detecting travellers infected with 2019-nCoV. *CMMID* **2020**. https://cmmid.github.io/ncov/airport_screening_report/airport_screening_preprint_2020_01_28.pdf.
34. Wearing, H.J.; Rohani, P.; Keeling M. J. Appropriate Models for the Management of Infectious Diseases. *PLOS Medicine* **2005** *2(7)*. https://doi.org/10.1371/journal.pmed.0020174.
35. Riley, S. *et al.* Transmission Dynamics of the Etiological Agent of SARS in Hong Kong: Impact of Public Health Interventions. *Science* **2003** *300(5627)*, 1961–1966. https://doi.org/10.1126/science.1086478.
36. Lauer S. *et al.* Update January 31: Modeling the Spreading Risk of 2019-nCoV. *JHU IDD* **2020**. https://systems.jhu.edu/research/public-health/ncov-model-2.
37. Nishiura, H. *et al.* The Rate of Underascertainment of Novel Coronavirus (2019-nCoV) Infection: Estimation Using Japanese Passengers Data on Evacuation Flights. *J. Clin. Med.* **2020**, *9(2)*, 419. https://doi.org/10.3390/jcm9020419.
38. Bolyai Institute, University of Szeged. Risk assessment of novel coronavirus 2019-nCoV outbreaks outside China. *Github* **2020**. https://github.com/zsvizi/corona-virus-2020.