The effects of border control and quarantine measures on the spread of COVID-19

M. Pear Hossain[a,b], Alvin Junus[a], Xiaolin Zhu[c], Pengfei Jia[d], Tzai-Hung Wen[e], Dirk Pfeiffer[f], Hsiang-Yu Yuan[a,*]

[a] *Department of Biomedical Sciences, Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong, Hong Kong*
[b] *Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Bangladesh*
[c] *Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong*
[d] *Academic Information Center, China Academy of Urban Planning and Design, Beijing, China*
[e] *Department of Geography, National Taiwan University, Taiwan*
[f] *Centre for Applied One Health Research and Policy Advice, City University of Hong Kong, Hong Kong*

ABSTRACT

The rapid expansion of coronavirus disease 2019 (COVID-19) has been observed in many parts of the world. Many newly reported cases of COVID-19 during early outbreak phases have been associated with travel history from an epidemic region (identified as imported cases). For those cases without travel history, the risk of wider spreads through community contact is even higher. However, most population models assume a homogeneous infected population without considering that the imported and secondary cases contracted by the imported cases can pose different risks to community spread.

We have developed an "easy-to-use" mathematical framework extending from a meta-population model embedding city-to-city connections to stratify the dynamics of transmission waves caused by imported, secondary, and others from an outbreak source region when control measures are considered. Using the cumulative number of the secondary cases, we are able to determine the probability of community spread.

Using the top 10 visiting cities from Wuhan in China as an example, we first demonstrated that the arrival time and the dynamics of the outbreaks at these cities can be successfully predicted under the reproduction number $R_0 = 2.92$ and incubation period $\tau = 5.2$ days. Next, we showed that although control measures can gain extra 32.5 and 44.0 days in arrival time through an intensive border control measure and a shorter time to quarantine under a low R_0 (1.4), if the R_0 is higher (2.92), only 10 extra days can be gained for each of the same measures. This suggests the importance of lowering the incidence at source regions together with infectious disease control measures in susceptible regions. The study allows us to assess the effects of border control and quarantine measures on the emergence and global spread of COVID-19 in a fully connected world using the dynamics of the secondary cases.

1. Introduction

On 31 December 2019, the World Health Organization (WHO) was alerted to several cases of pneumonia infections in Wuhan City, Hubei Province of China (World Health Organization, 2020). The cause of the pneumonia was later identified as a novel coronavirus, referred to as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which is genetically closely related to the Middle Eastern Respiratory Syndrome virus (MERS-CoV) and the Severe Acute Respiratory Syndrome virus (SARSCoV) (World Health Organization, 2019). The novel virus was able to establish human-to-human transmission (Chan et al., 2020) and caused high mortality and morbidity rates. Just after one month, more than a hundred confirmed cases have been reported all over the world outside China, not only in nearby countries like Japan, Korea, Singapore, but also in Europe and the Americas (The Guardian, 2020). WHO has since declared the outbreak an international emergency.

During the outbreak emergence, one of the most urgent public health tasks is to prevent the spread of the virus from an epidemic source region to other regions within a country or globally. Because a person who was infected can travel to another region and spread the virus, COVID-19 continued to pose a severe threat to other regions through transportation services. Many newly reported cases of this new coronavirus infection in other cities or countries before community spread have been associated with travel history from an epidemic source region or contact history to people from the region, referred as the imported cases and the secondary cases transmitted from the imported cases respectively. Once the secondary cases continue to transmit to more local cases, infection chain is established in the community, community spread begins subsequently (CDC, 2020).

The importance of prediction of infectious diseases based on transportation network information has already been highlighted in many previous studies (Hwang et al., 2012; Nicolaides et al., 2012). The

---

probability of emergence and the arrival time of the emergence can be estimated using different approaches through model simulation with intensive computation or analytical expressions based on complex network structure with Poisson processes (Gautreau et al., 2008; Tomba and Wallinga, 2008; Wang and Wu, 2018; Ruben et al., 2010). However, due to intensive simulation or complex structure, these studies may either provide limited insights into public health or may be difficult to be implemented. A swift response in public health decisions is required for society's infectious disease emergency preparedness. From infectious disease control perspective, the major questions here are whether we can i) assess the risk for COVID-19 to spread to other cities or countries to cause community spread; and ii) evaluate the effects of infectious disease control by cessation of population movement (e.g., lockdown, border control, or quarantine measures) on the outbreak spreading. Thus, an easy-to-use mathematical formula that can be embedded in a classical meta-population model to evaluate the risk of community spread and the effects of control measures is needed.

The increasing number of secondary cases indicate a significant risk of community spread. Using the Severe Acute Respiratory Syndrome (SARS) outbreak as an example, the progression of the SARS outbreak during a given time was generally following this order. First, people migrated from a city with the outbreak. Thus, infected cases arrived in other cities or countries as imported cases; second, the imported cases began to infect those who had very close contacts with them, resulting in secondary cases (Liu et al., 2020a,b); third, local transmission began after infections occurred for people without contact history with the imported cases or travel history to the outbreak source (Riley et al., 2003). After these processes occurred repeatedly in many regions, SARS eventually spread rapidly throughout the world via air travel. Normally the imported cases can be soon isolated or quarantined due to the strict control policy against COVID-19 spread. However, contact tracing and isolation of cases are more difficult to be performed on secondary cases than imported, especially when R_0 is high (Hellewell et al., 2020). To prevent community transmission to happen, early prediction of the cumulative secondary cases generated by the imported cases is extremely critical.

Although disease compartment models with meta-population are widely used with migration or transportation data, the complexity of using these models sometimes infuriate infectious disease modelers when an urgent response is needed because most of the meta-population models do not offer a simple way to estimate the dynamics of imported and secondary cases (Riley et al., 2003; Wang and Wu, 2018). How to produce the dynamics of these cases in a simple way when the incidence is still increasing at the source remains to be answered. In addition, in order to evaluate the impact of infectious disease control measures on virus spread, there is a need to estimate the probability of emergence and the arrival time of the emergence through secondary infections under different control measures.

In this study, we have built a meta-population model based on a classical SIR model coupled with a mobility matrix and mathematical expressions to understand the outbreak spreading dynamics stratified by imported, secondary and other local cases at different cities. We have generate mathematical expressions that can be embedded into the SIR model to estimate the outbreak potential at neighboring cities. These formulas allowed us to calculate the dynamics of the first wave (imported cases) and the second wave of transmission (secondary cases produced by the imported cases). Using the cumulative number of the secondary cases, we thus predicted the probability of outbreak emergence and evaluated border closure and quarantine measures against this nation-wide outbreak spreading. Gain time before outbreak emergence was predicted under different control measures with different R_0 settings.

2. Materials & Methods

2.1. Meta-population model

Assuming the newly emergence of COVID-19 causes an outbreak at location i, during the emergence, the changes of the numbers of infectious cases I_j at a different location j can be determined using a simple Susceptible, Infected, and Recovered (SIR) meta-population model with a mobility matrix (contact mixing at the population level):

$$\frac{dI_i}{dt} = \beta M_{ii}\frac{S_i}{N}I_i - M_{ji}I_i - \frac{1}{T_g}I_i \tag{1}$$

$$\frac{dI_j}{dt} = M_{ji}I_i + \beta M_{jj}\frac{S_j}{N}I_j - \frac{1}{T_g}I_j \tag{2}$$

where β is the baseline transmission rate that can be estimated from $\frac{R_0}{T_g}$, R_0 is the basic reproduction number, T_g is the generation time, M_{ii} is the human mobility rate within the source location i, M_{ji} is the mobility rate from i to j, and I_j is the number of infected individuals at the location j. Our aim was to develop a meta-population model embedded with analytical expressions that can stratify the imported cases and the secondary cases produced by the imported cases along with other infected individuals with border control and quarantine measures. We modified this simple meta-population model by introducing the effect of border control and the quarantine. The mobility rate was multiplied by $(1 - c)$, where c represents a border control measure. When c is higher up to 1, the mobility rate is reduced to zero. The infected cases were quarantined on average T_{qr} days after they are transmitted. After derivation (steps are described in later sections), the final model became:

$$\frac{dI_i}{dt} = \beta M_{ii}\frac{S_i}{N}I_i - \beta(1 - c)M_{ji}\eta(k = 0, .)I_i \tag{3}$$

$$\frac{dI_j}{dt} = \beta M_{jj}\frac{S_i}{N}Sec + \beta M_{jj}\frac{S_j}{N}I_j - \frac{1}{T_g}I_j \tag{4}$$

$$\frac{dImp}{dt} = \beta(1 - c)M_{ji}\eta(k = 0)I_i - \left(\frac{1}{T_g} + \frac{1}{T_{qr}}\right)Imp \tag{5}$$

$$\frac{dSec}{dt} = \beta(1 - c)M_{ji}\eta(k = 1)I_i - \left(\frac{1}{T_g} + \frac{1}{T_{qr}}\right)Sec \tag{6}$$

where Imp and Sec represent the imported and secondary cases produced by the imported cases (we will use secondary cases to denote this group in the remaining parts). We introduced an $\eta(k); k = 0, a, \ldots$ function to map the number of the infected at the source i to the imported and secondary cases at j given different border control and quarantine measures. It is a function of four epidemiological parameters such as the basic reproduction number R_0, the generation time T_g, time to quarantine T_{qr} and incubation period τ. The term $\eta(k = 0)$ calculated the changes during first wave transmission (imported) and $\eta(k = 1)$ calculated the changes during second wave transmission (secondary infected cases produced by the imported cases) under quarantine.

During the early outbreak phase, because the susceptible population S was so close to the population size N, therefore, we assumed $\frac{S}{N} \approx 1$. Because M_{ii} and M_{jj} are both near one (every day, more than 99.99% of individuals stay in the same location), we thus ignored the variables M_{ii} and M_{jj} to increase readability in the remaining sections.

2.2. Calculating the arrival rate of imported cases

In order to calculate the imported cases, we assumed that infected cases could pass the border screening or move to another location only during their incubation period (referred to as exposed cases). We calculated the number of exposed cases at time s by including an

incubation period τ. Therefore, the number of cases I_i in Eq(1) that were within the incubation period at a specific time s were $I_{i\_e\,xposed}(s) = \int_0^\infty \beta I_i(s-a)(1-F(a))\mathrm{d}a$, where $F(a) = 1 - e^{-(\frac{1}{\tau}+\frac{1}{T_g})a}$ is the cumulative distribution function of the exposed cases that were transmitted a days ago but before recovery. The longer an incubation period was, the more total imported cases were produced in a given duration. After replacing the number of infected cases by the number of exposed cases, we obtained the rate of imported cases:

$$\begin{aligned}\mathrm{Imp}^+ &= (1-c)M_{ji}I_{i\_e\,xposed} \\ &= \frac{R_0}{T_g}\eta(k=0)(1-c)M_{ji}I_0 e^{\Lambda t}\end{aligned} \tag{7}$$

where Λ is the growth rate and can be calculated as $\Lambda = \frac{R_0-1}{T_g}$ and $\eta(k) = \left(\frac{R_0 T_{qr}}{R_0 T_{qr}+T_g}\right)^k\left(\Lambda + \frac{1}{\tau} + \frac{1}{T_g}\right)^{-1}$. If all the infected cases can move to a different place as the incubation period τ is quite long enough, the formula is reduced to $(1-c)M_{ji}I(t)$.

2.3. Calculating numbers of imported and secondary cases

To calculate the number imported (Imp) and secondary cases (Sec), we had the following formulas now:

$$\frac{\mathrm{d}Imp}{\mathrm{d}t} = Imp(t)^+ - \left(\frac{1}{T_g}+\frac{1}{T_{qr}}\right)Imp \tag{8}$$

$$\frac{\mathrm{d}Sec}{\mathrm{d}t} = \beta\frac{S_i}{N}Imp(t) - \left(\frac{1}{T_g}+\frac{1}{T_{qr}}\right)Sec \tag{9}$$

where T_{qr} is the time to quarantine. The number of cumulative imported cases can thus be derived during a certain period of time when the incidence is still exponentially increasing at the source location. Assuming $\frac{S_i(t)}{N} \approx 1$, after solving the differential equation (detail derivation is available in Eq. S4 to S11), the numbers of the imported cases and the secondary cases at time t were:

$$Imp(t) = \eta(k=1)(1-c)M_{ji}I_0 e^{\Lambda t} \tag{10}$$

$$Sec(t) = \eta(k=2)(1-c)M_{ji}I_0 e^{\Lambda t} \tag{11}$$

Consequently, the cumulative number of imported and secondary cases at time t under incubation period were:

$$CI_{Imp}(t) = \frac{R_0}{T_g}\eta(k=0)(1-c)M_{ji}I_0\left(\frac{e^{\Lambda t}-1}{\Lambda}\right) \tag{12}$$

$$CI_{Sec}(t) = \frac{R_0}{T_g}\eta(k=1)(1-c)M_{ji}I_0\left(\frac{e^{\Lambda t}-1}{\Lambda}\right) \tag{13}$$

Because before the outbreak occurred (or before community spread) at a location j, infections only happened when transmission events occurred between the imported cases and the susceptible individuals at j (we call it the second wave of transmission), it is essential to estimate the cumulative number of the imported cases (CI_{Imp}; the first wave of transmission) and the cumulative secondary cases transmitted from those imported cases (CI_{Sec}; the second wave of transmission). We were able to deduce tertiary cases but the number will be relatively small comparing to the imported and the secondary.

Thus, we have a simple formula to predict certain important infection numbers before an outbreak emerges in a specific location using four epidemiological parameters including the basic reproduction number, generation time, incubation period and time to disease detection after onset along with a contact matrix. This framework provides a more generalized expression to estimate the cumulative imported cases (first wave of transmission) and the cumulative secondary infected cases generated by the imported cases (second wave of transmission) at connected cities or locations during a certain period of time when the incidence is exponentially increasing at the source.

2.4. Constructing contact matrix

Airline passenger data were collected from the International Air Transport Association (IATA) database. We collected the actual passenger data for the top 10 visiting cities leaving from Wuhan Tianhe International Airport before the lockdown of Wuhan city from 30 December to 20 January, 2020. Note that we did not have data for railroad and other forms of transport, and thus made an assumption that the total number of travelers is 4 times higher than that of air transport except for certain cities on Hainan island that have no road connections to Wuhan. We made this assumption because that the number of train passengers is few times higher than that of airline in China (Statista, 2018) and the results from a population migration database suggested a similar ratio (The Tencent database, 2020). The number of daily passengers between different cities were used to generate the mobility rate M_{ji} between locations i and j. For example, we divided a daily passenger number by total population size in Wuhan to represent the contact rate between Wuhan city and any other connected city j. We denote $i = 0$ as the index for the source city. Therefore, $M_{00} = 1$ is the base level of contact rate within the source city. j is a number to represent the index of the top visiting cities from first to last following the order in the list. Because the estimated number of people leaving from Wuhan to the first visiting city (Beijing) is 10793.5 per day, given that the population size of Wuhan is about 11 million, the daily percentage of people leaving to the 1st visiting city can be approximated to 0.001. Thus, we have $M_{10} = 0.001$. We used the same approach to determine M_{j0} for $j = 2, 3, \ldots, 10$.

2.5. Determining outbreak potential

Next we considered the outbreak potential, defined as the probability of outbreak emergence given the number of cumulative cases. At the initial stage, if there were n cumulative infected cases, the chance of the viruses to cause an outbreak is $p_{j,n} = 1 - \frac{1}{R_{j\_i\,ni}^n}$ (Allen and Lahodny, 2012). From the observed number of cumulative cases in the top 10 visiting cities, we have found that in general, after the number was reaching or above 8, the number would begin to increase sharply except Nanning. Therefore, we determined a critical threshold number ν to be 8 and set $p_{j,\nu} = 50\%$ after we compared the trends of the top 10 visiting cities. 50% of the cities demonstrated rapid growth of the numbers of infected cases once the numbers reached or near the threshold. We thus obtained the effective reproduction number $R_{j\_i\,ni} = 1.0905$ to represent the transmissibility during the second wave of transmission at location j. Note that this number represents the epidemic growth under control measures before community spread. $R_{j\_i\,ni}$ indicates the average number of transmissions that are generated from the secondary cases (Sec) before the community spread. Because the nation-wide alert has already been received at different cities after 31 December, 2019, and many infectious disease control measures have been implemented, $R_{j\_i\,ni}$ was expected to be lower than R_0. Given the $R_{j\_i\,ni}$ number, we used the cumulative number of secondary cases CI_{Sec} at the location j to calculate the probability of outbreak emergence as $p_{j,CI_j} = 1 - \frac{1}{R_{j\_i\,ni}^{CI_j}}$ (to simplify the notation, the CI_j is used to represent CI_{Sec}). We defined the critical arrival time such that the probability of outbreak emergence p_{j,CI_j} was larger than 50%.

3. Results

3.1. Model of outbreak spreading

The cities that most Wuhan citizens moved to in the first month after the outbreak of the COVID-19 was reported in Wuhan posed a higher potential of outbreak emergence. We collected airline passenger data during 22 days after the outbreak emerged between 30 December to 20 January, before the lockdown of Wuhan was implemented on 23
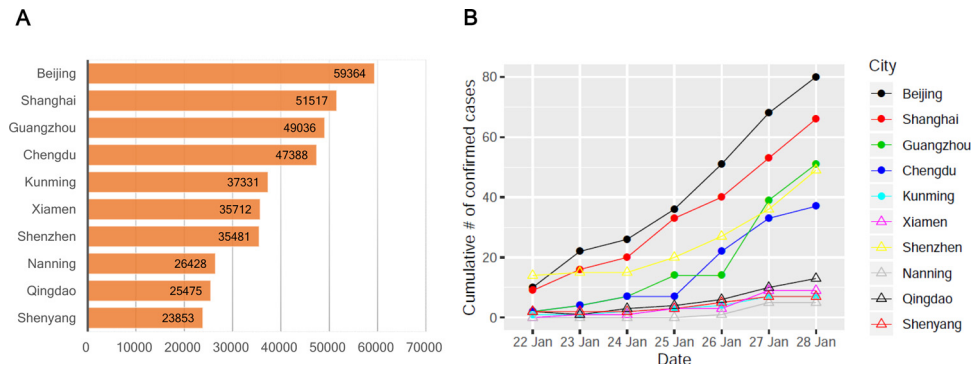
Fig. 1. Statistics of airline passenger from Wuhan Tianhe International Airport and COVID-19 confirmed cases. (A) Number of airline passengers from Wuhan to top 10 visiting cities between 30 December, 2019 to 20 January, 2020. (B) Number of confirmed cases at top 10 visiting cities between 22 January to 28 January, 2020.

January, 2020. The daily airline passenger number from Wuhan to the top 10 visiting cities were extracted with the 1779.9 persons on average (Fig. 1A). We estimated the total passengers leaving out from Wuhan using a human migration map airline passengers. We obtained the estimated numbers of total passengers given the proportion of airline passenger was about 23% of all transportation in early January (The Tencent database, 2020). An increasing number of the confirmed cases was observed among the top visiting cities, including Beijing, Shanghai, and Guangzhou (Fig. 1B).

The community spread began as the number of the imported cases and the associated secondary cases generated by the imported cases accumulated to a certain number (Fig. 2A). Thus, the increasing number of the imported cases can be correlated to the outbreak spreads through the number of departure passenger data. We developed a mathematical framework deriving from a standard meta-population model coupling with a migration matrix and incubation period with different control measures (Fig. 2B) to calculate the number of the total imported cases and the cumulative secondary cases.

3.2. Outbreak spreads in Mainland China

Outbreak spreads in Mainland China were reconstructed and compared with different transmissibilities and incubation periods. We first calculated the cumulative number of secondary infected cases produced by imported cases among top 10 visiting cities from Wuhan under a scenario corresponding to an R_0 of 2.92 (Liu et al., 2020a,b) (the analyses for other scenarios such as mild R_0(1.68) and low R_0(1.4) are given in Figure S1 and Figure S2), a generation time of 8.4 days (Lipsitch et al., 2003), and an incubation period of 5.2 days for the outbreak in Wuhan city. The migration matrix M was constructed using the average of the airline passengers data during 22 days in January

among these cities (*materials and methods*). The cumulative number of secondary infected individuals generated by the imported cases moving from Wuhan was calculated (Fig. 3A). A threshold number of cases $\nu = 8$ was used to indicate a higher than 50% probability of community spread will occur if the cumulative number of secondary cases is over that threshold.

The arrival times of outbreak emergence (defined as the time to reach above this threshold) for the top 10 cities were between 16 - 20 days (Fig. 3A), with the average arrival time 18 days (corresponding to 18 January). On day 28, the secondary cases rise to 139 persons for the top city Beijing, which makes the total number of secondary cases among all the top 10 visiting cities to 915.7 persons. The outbreak potentials of the cities were assessed on 18 January. 7 out of 10 cities have a probability of outbreak emergence larger than 50% (Fig. 3B and Figure S6). Among those 7 cities, 5 of them had a high number of actual confirmed cases more than 36 on 28 January (10 days later), only 3 of them maintained low case numbers below 10. Taking into account that the actual lag of reporting time was about 10 days, the probability of outbreak emergence on 18 January indicates the level of the outbreak potential well (Fig. 1B).

The predicted reporting delay was very close to the actual reporting lag. The average actual lag was calculated to be 10.30 days after counting the difference between the average date of onset peak (among 8 days with highest number of cases) and the average date of diagnosis peak (among 8 days with highest number of cases) shown in the recent report of 72314 cases from the Chinese Center for Disease Control and Prevention (Wu and McGoogan, 2020). Note that we did not consider the long tail because no full longitudinal data were available yet. The estimated value of the average lag time using maximum likelihood approach based on a binomial distribution of the infected cases in different cities to fit 10 cities together was 10.52 days (10.27 − 10.78;
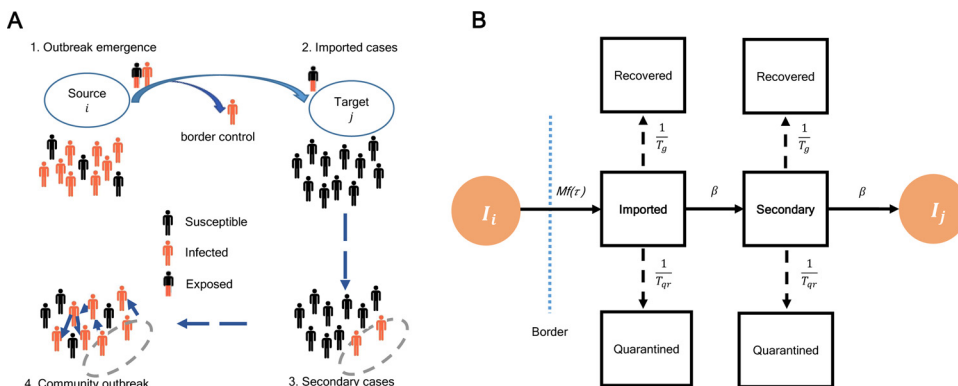

Fig. 2. COVID-19 outbreak spread from source to community transmission. (A) Outbreak progression from source to community spread. The imported cases arrive after passengers passed the border control. The secondary cases produced by the imported cases eventually cause the community outbreak. The community outbreak starts after *t*. (B) Mathematical model framework for COVID-19 estimates of secondary cases. I_i and I_j represent the numbers of infected cases in a source location *i* and a target location *j* respectively. *M* is the mobility rate and $f(\tau)$ is a function of incubation period that represent the percentage of infected cases that can pass the border (dashed line). T_g and T_{qr} are the generation time and time to quarantine respectively. β estimates the ratio $\frac{R_0}{T_g}$, where R_0 is the basic reproduction number.
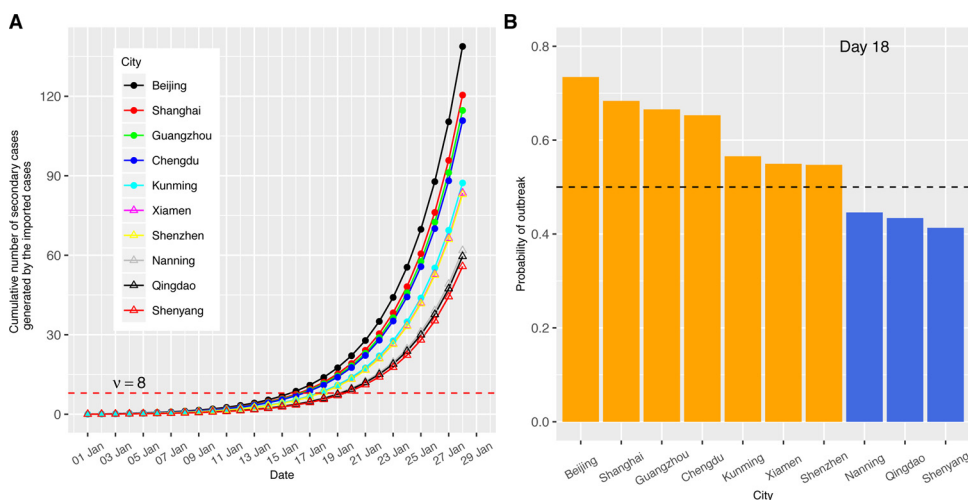
A

B

95% confidence interval) for incubation period 5.2 days (Figure S3). Note that when we fitted each city individually, the reporting delay ranges from 8.0 to 18.3 days (Figure S4), with an average of 12.4 days. When only the top five cities with highest confirmed cases are considered, the reporting delay ranges from 8.0 to 11.2 days with an averages at 9.2 days. We did not rule out the possibility that the long delay time, such as 18.3 days, was due to actual delay of the outbreak emergence. We further investigated whether any over dispersion effect exists on reporting delay using a negative binomial distribution (Table S2). The results were then compared with the binomial model. We found that both models produced similar estimates of reporting delay for each of the top 10 cities. The average reporting delay was 12.6 days when over dispersion was included. Overall, the predicted cumulative number of both imported and secondary cases after adjusted by the reporting lag time of top 10 visiting cities demonstrated a similar increasing trend in cumulative numbers of confirmed cases during each early emergence period (Fig. 4 and Figure S5).

In contrast, a longer incubation period shortened the outbreak arrival time by allowing more secondary cases. Considering the same R_0 setting but with an incubation period of 14 days, on day 28 the total number of secondary cases raised to 1179.8 persons for all the top 10 visiting cities in which Beijing contributes 180 persons in the total secondary cases. Therefore, the longer incubation period allowed 22% more secondary cases. The arrival time of outbreak emergence were 14-18 days (Fig. 5A), which were 2 days earlier than using 5.2 days incubation period. On day 18, all top 10 cities except Shenyang have outbreak probability more than 50% (Fig. 5B). The results confirmed that if the incubation period is long, more ill people can move to other cities.

For each of the top 10 cities, the arrival times of outbreak emergence for both 5.2 and 14 days incubation periods were determined. The actual order of the arrival time for the top four cities, including Beijing, Shanghai, Guangzhou and Chengdu were successfully predicted (Table 1). A longer incubation period produced a shorter arrival time.

3.3. Impact assessment of border control and quarantine

We first evaluated the effects of border control measures by changing the control rate c to reduce the transportation between two cities under three different R_0 settings. Border control or other special events, like lunar new year, can affect the traveling rates. Ideally, complete cessation of population movement between cities or isolation of every susceptible case from the source city can reduce transmission events most however it happens rarely. Then the extra arrival gain times of outbreak emergence comparing to a baseline setting using Beijing city were calculated.

Under a low R_0 (1.4), the border control measure that reduced 90% of the passenger numbers gained extra 32.5 days of outbreak arrival time (Fig. 6A). Under a medium R_0 (1.68), we found that the effect of border control was weaker but still created 20.0 extra days under the same control level. However, under the high R_0 (2.92), the effect on reducing the chance of outbreak emergence was very low, with only 10 extra days obtained.

Next, we calculated gain time by time to quarantine under different R_0 settings. Under the low R_0 (1.4), if the individual was quarantined immediately (happened in one day after the person became infectious), the gain time became as large as 44.0 days (Fig. 6B). Under the medium R_0 (1.68), we found that the quarantine effect was approximately half of the low R_0 scenario (24.1 days) using the same time to quarantine. However, under the high R_0 (2.92), the effect of quarantine was much lower, with only 10.0 days was gained.

4. Discussion

The current COVID-19 epidemic marks the third time in 20 years that a member of the family of coronaviruses (CoVs) has caused an epidemic employing its zoonotic potential, for example, from bats (Zhou et al., 2020). SARS-CoV-2 is able to establish between human-to-human transmission (Chan et al., 2020) and has been currently spreading from Wuhan to many nearby cities and countries. It is hypothesized that the rate of transmission between different cities or countries is related to the number of people moving from different locations to Wuhan. About two weeks later after 31 December, using the number of cases detected outside China, it has been inferred that more than a thousand individuals (with an estimated mean 1723) had had an onset of symptoms by 12 January, 2020 (Imai et al., 2020).

The study demonstrated that after the reporting delay was estimated, the dynamics of the outbreaks at connected cities can be successfully reconstructed using both the imported and the secondary cases. The result implies that all the connected (direct or indirect) countries are having a great risk of outbreak. The question is about when the outbreak will arrive (CNN, 2020) and how to delay the arrival time to have a better preparation. The challenge, however, is that we lack a simple and accurate tool for assessing outbreak emergence risk and subsequently the required levels of border control and quarantine measures to prevent additional outbreaks.

Until recently, although some studies have been done to predict the spreading of this new disease using air and other forms of transport information (Hwang et al., 2012; Nicolaides et al., 2012; Wu et al., 2020; Lai et al., 2020; Cao et al., 2020), none of studies were designed to estimate the dynamics of the imported and secondary cases. The benefit of stratifying the imported and secondary cases in a disease
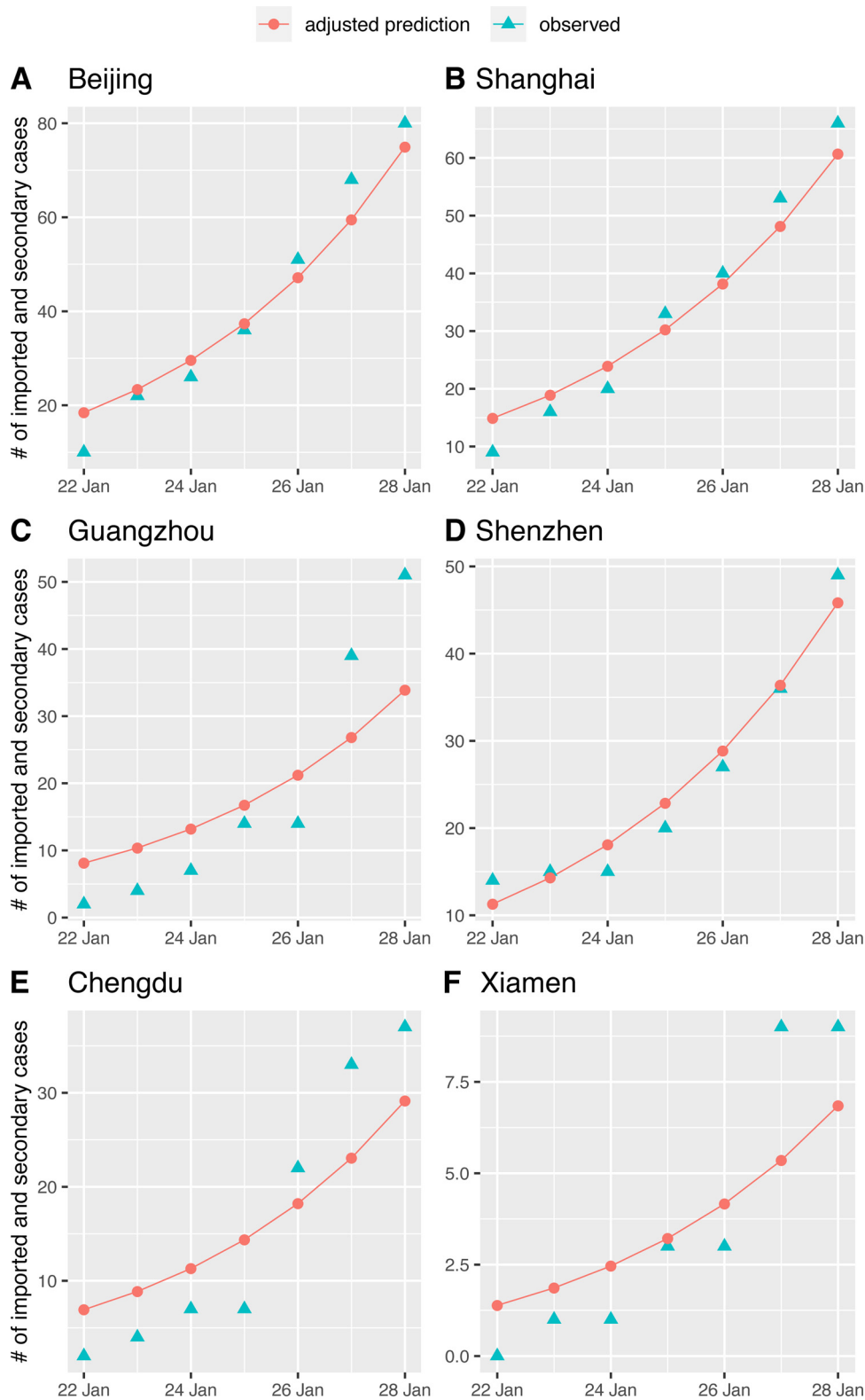
Fig. 4. Observed number of confirmed cases and predicted number of imported and secondary cases. The predicted number of cases were adjusted by the reporting delay after using maximum likelihood estimation. The six cities that have the actual earliest arrival times are listed (four other remaining cities are given in Figure S5).

transmission model is to provide a risk assessment of community spread. Because most of the imported cases can be detected more easily under 14 days quarantine from the passengers coming from the epidemic source region, thus the risk of outbreak is not primarily linked to the number of the imported cases. However, secondary cases, without travel history to the epidemic source region, are more difficult to be identified or quarantined before disease onset and thus are more easily to become undetected cases in a community.

We evaluated the effectiveness of different infectious disease control on the secondary cases in order to estimate the arrival time of future
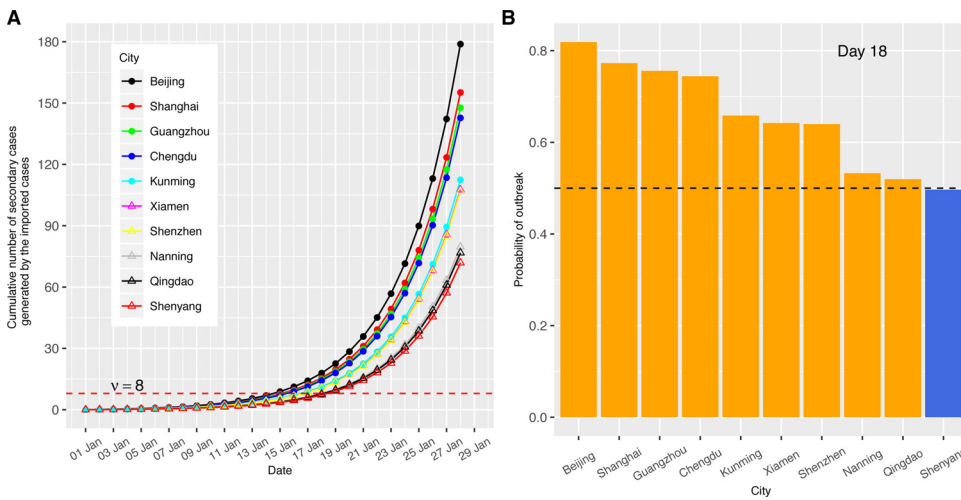
Fig. 5. Outbreak potential estimated from the secondary cases contacted by imported cases. A higher $R_0 = 2.92$ scenario with incubation period $\tau = 14$ days and time to quarantine $T_{qr} = 2$ days were used. (A) Number of cumulative secondary cases generated by imported cases. The secondary infections are listed among the top 10 visiting cities from Wuhan. $\nu = 8$ is the critical threshold number; (B) Probability of outbreak emergence in different cities at critical time (18 days).

community spread. Recent studies have shown the importance of modeling in infectious disease control (Fraser et al., 2004). A recent study has used a modeling approach to forecast the dynamics of outbreak spreading (Wu et al., 2020). We developed an "easy-to-use" mathematical formula that are able to have an analytical solution of the first wave transmission (imported cases) and the second wave transmission (secondary cases generated by the imported cases). With these numbers, we are able to evaluate the impact of border control and quarantine measures. Surprisingly, under the higher R_0 setting (2.92), the effect on obtaining 10 extra days requires an enhanced border control measure to reduce more than 90% of the passengers or a very efficient quarantine measure. The results suggest that if the epidemic growth at the source location is high, even a near full-scale border control without proper quarantine measures, will have only limited effects. The transmission waves can be treated as a branching process. However, instead of using the offspring variability to estimate the probability of extinction, we adopted a classical way to derive the probability of extinction that was based on R_0 or effective R.

We have learnt from the previous SARS outbreak that it is crucial to implement rapid infection control measures to limit the impact of epidemics, both in terms of preventing more casualties and shortening the epidemic period. Delaying the implementation of control measures by 1 week would have nearly tripled the epidemic size and would have increased the expected epidemic duration by 4 weeks (Wallinga and Teunis, 2004). Previous studies showed that control measures at international cross-borders and screening at borders are influential in

mitigating the spread of infectious diseases (Wang et al., 2019) (Priest et al., 2015). Cross-boarder screening system to prevent infectious disease outbreak is important but cannot successfully prevent ill persons from arriving during their incubation period.

Full-scale border control measures to prevent the spread of the virus have been discussed in many countries in the world. At the same time, a lockdown of Wuhan city (border control from leaving out) has already been imposed. Here the model we constructed can be used to estimate the dynamics of imported and secondary cases using transportation data with different control measurements. The framework can be extended to multiple infected sources to multiple target cities without increasing the complexity of the computation dramatically. Hence, the model proposed in the current study could provide a risk assessment of COVID-19 global spreading in a highly connected world.

Declaration of Interests

All authors declare no competing interests.

Author Statement

M. Pear Hossain: Methodology, Software, Writing - Original Draft. Alvin Junus: Methodology, Software. Xiaolin Zhu: Conceptualization. Pengfei Jia: Investigation. Tzai-Hung Wen: Conceptualization. Dirk Pfeiffer: Conceptualization. Hsiang-Yu Yuan: Conceptualization, Methodology, Supervision, Writing - Original Draft.

Table 1
Actual and predicted arrival time of outbreak emergence at top 10 connected cities in China. $R_0 = 2.92$ with incubation time $\tau = 5.2$ days and $\tau = 14$ days were used. CI_{Sec} is the cumulative number of secondary infected cases generated by the imported cases. The actual arrival time of outbreak is defined as the the date when the number of cumulative cases is larger than the threshold number 8 and the number of newly reported cases is larger than 5. 9.2 days reporting lag between the date of onset and the date of diagnosis was estimated using the top five cities with most number of confirmed cases.

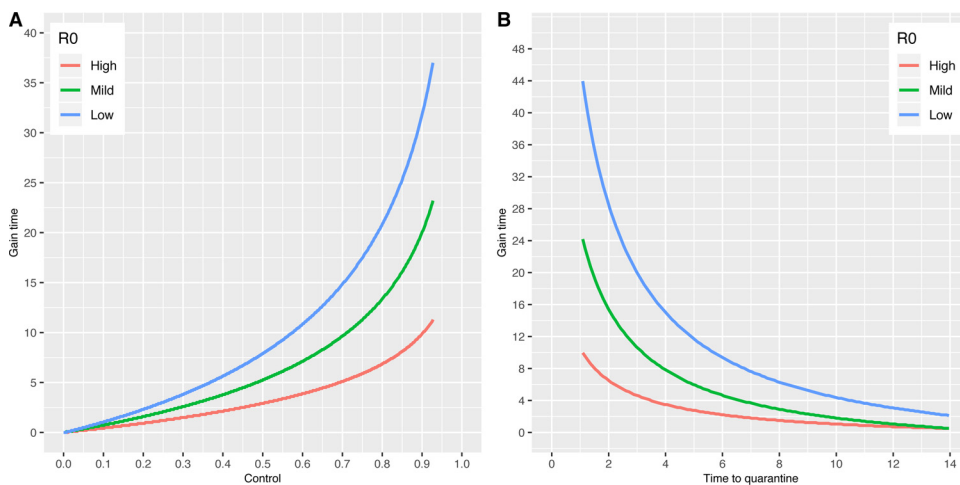| City	Actual arrival time (day)	$\tau = 5.2$			$\tau = 14$		
		Predicted arrival time (day)	CI_{Sec} at critical time (day 18)	Probability of outbreak at critical time (day 18)	Predicted arrival time (day)	CI_{Sec} at critical time (day 18)	Probability of outbreak at critical time (day 18)
Beijing	< 23	16	13.9	0.734	15	17.9	0.819
Shanghai	< 23	17	12.1	0.684	16	15.6	0.773
Guangzhou	25	17	11.5	0.666	16	14.8	0.756
Chengdu	26	17	11.1	0.653	16	14.3	0.744
Kunming	29	18	8.7	0.566	17	11.3	0.658
Xiamen	27	18	8.4	0.550	17	10.8	0.642
Shenzhen	25	18	8.3	0.547	17	10.7	0.640
Nanning	> 30	20	6.2	0.446	19	8.0	0.533
Qingdao	> 30	20	6.0	0.434	19	7.7	0.520
Shenyang	> 30	20	5.6	0.413	19	7.2	0.497

Fig. 6. Assessment of border control and quarantine effects on outbreak arrivals. (A) Gain time of outbreak emergence by the rates of successful border control. The effects of border control on gain time under a low R_0 (1.4), mild R_0 (1.68), high R_0 (2.92) were plotted in blue, green, and red colors. (B) Gain time of outbreak emergence by time to quarantined. The effects of border control on gain time under a low R_0 (1.4), mild R_0 (1.68), high R_0 (2.92) were plotted using the same color codes as A. The passenger data of the top visiting city Beijing was used to generate the baseline arrival time. To get the gain time, the arrival time using different infectious disease control measures was calculated and was subtracted by the baseline arrival time.

Acknowledgments

Appendix A. Supplementary Data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.epidem.2020.100397.

References

World Health Organization (WHO), 2020. Pneumonia of unknown cause - China. https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/.

World Health Organization (WHO). Coronavirus disease (COVID-19) outbreak. https://www.who.int/emergencies/diseases/novel-coronavirus-2019.

Chan, Jasper Fuk-Woo, Yuan, Shuofeng, Kok, Kin-Hang, To, Kelvin Kai-Wang, Chu, Hin, Yang, Jin, Xing, Fanfan, Liu, Jieling, Yip, Cyril Chik-Yan, Poon, Rosana Wing-Shan, et al., 2020. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. The Lancet 6736 (20), 1–10.

The Guardian, 2020. Coronavirus map: how Covid-19 is spreading across the world. https://www.theguardian.com/world/ng-interactive/2020/feb/26/coronavirus-map-how-covid-19-is-spreading-across-the-world.

Center for Disease Control and Prevention (CDC), 2020. Coronavirus in Japan. https://wwwnc.cdc.gov/travel/notices/alert/coronavirus-japan.

Hwang, Grace M., Mahoney, Paula J., James, John H., Lin, Gene C., Berro, Andre D., Keybl, Meredith A., Michael Goedecke, D., Mathieu, Jennifer J., Wilson, Todd, 2012. A model-based tool to predict the propagation of infectious disease via airports. Travel Medicine and Infectious Disease 10 (1), 32–42.

Nicolaides, Christos, Cueto-Felgueroso, Luis, González, Marta C., Juanes, Ruben, 2012. A metric of influential spreading during contagion dynamics through the air transportation network. PLoS ONE 7 (7).

Gautreau, Aurélien, Barrat, Alain, Barthelemy, Marc, 2008. Global disease spread: statistics and estimation of arrival times. Journal of theoretical biology 251 (3), 509–522.

Tomba, Gianpaolo Scalia, Wallinga, Jacco, 2008. A simple explanation for the low impact of border control as a countermeasure to the spread of an infectious disease. Mathematical biosciences 214 (1-2), 70–72.

Wang, Lin, Wu, Joseph T., 2018. Characterizing the dynamics underlying global spread of epidemics. Nature Communications 9 (1), 1–11.

Kubiak, Ruben J., Arinaminpathy, Nimalan, Mclean, Angela R., 2010. Insights into the evolution and emergence of a novel infectious disease. PLoS Computational Biology 6 (9).

Liu, Yang, Eggo, Rosalind M., Kucharski, Adam J., 2020a. Secondary attack rate and superspreading events for SARS-CoV-2. The Lancet [In Press].

Riley, Steven, Fraser, Christophe, Donnelly, Christl A., Ghani, Azra C., Abu-Raddad, Laith J., Hedley, Anthony J., Leung, Gabriel M., Ho, Lai-Ming, Lam, Tai-Hing, Thach, Thuan Q., et al., 2003. Transmission dynamics of the etiological agent of sars in hong kong: impact of public health interventions. Science 300 (5627), 1961–1966.

Hellewell, Joel, Abbott, Sam, Gimma, Amy, Bosse, Nikos I., Jarvis, Christopher I., Russell, Timothy W., Munday, James D., Kucharski, Adam J., John Edmunds, W., Sun, Fiona, Flasche, Stefan, Quilty, Billy J., Davies, Nicholas, Liu, Yang, Clifford, Samuel, Klepac, Petra, Jit, Mark, Diamond, Charlie, Gibbs, Hamish, van Zandvoort, Kevin, Funk, Sebastian, Eggo, Rosalind M., 2020. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. The Lancet Global Health [ONLINE FIRST].

Statista. Number of train passengers in China from 2008 to 2018. https://www.statista.com/statistics/276054/number-of-train-passengers-in-china/.

The Tencent database. Population migration map. https://heat.qq.com/qianxi.php.

Allen, Linda J.S., Lahodny Jr., Glenn E., 2012. Extinction thresholds in deterministic and stochastic epidemic models. Journal of Biological Dynamics 6 (2), 590–611.

Liu, Tao, Hu, Jianxiong, Kang, Min, Lin, Lifeng, Zhong, Haojie, Xiao, Jianpeng, He, Guanhao, Song, Tie, Huang, Qiong, Rong, Zuhua, et al., 2020b. Transmission dynamics of 2019 novel coronavirus (2019-nCoV). bioRxiv.

Lipsitch, Marc, Cohen, Ted, Cooper, Ben, Robins, James M., Ma, Stefan, James, Lyn, Gopalakrishna, Gowri, Chew, Suok Kai, Tan, Chorh Chuan, Samore, Matthew H., et al., 2003. Transmission dynamics and control of severe acute respiratory syndrome. Science 300 (5627), 1966–1970.

Wu, Zunyou, McGoogan, Jennifer M., 2020. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72?.314 Cases From the Chinese Center for Disease Control and Prevention. JAMA [New Online].

Zhou, Peng, Yang, Xing-Lou, Wang, Xian-Guang, Hu, Ben, Zhang, Lei, Zhang, Wei, Si, Hao-Rui, Zhu, Yan, Li, Bei, Huang, Chao-Lin, et al., 2020. Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. bioRxiv.

Imai, Natsuko, Dorigatti, Ilaria, Cori, Anne, Riley, Steven, Ferguson, Neil M., 2020. Estimating the potential total number of novel coronavirus cases in wuhan city, china. https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/2019-nCoV-outbreak-report-17-01-2020.pdf.

Cable News Network (CNN), 2020. CDC official warns Americans it's not a question of if coronavirus will spread, but when. https://edition.cnn.com/2020/02/25/health/coronavirus-us-american-cases/index.html.

Wu, Joseph T., Leung, Kathy, Leung, Gabriel M., 2020. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet 395 (10225), 689–697.

Lai, Shengjie, Bogoch, Isaac, Ruktanonchai, Nick, Watts, Alexander, Li, Yu, Yu, Jianzing, Lv, Xin, Yang, Weizhong, Yu, Hongjie, Khan, Kamran, et al., 2020. Assessing spread risk of wuhan novel coronavirus within and beyond china, january-april 2020: a travel network-based modelling study. medRxiv.

Cao, Zhidong, Zhang, Qingpeng, Lu, Xin, Pfeiffer, Dirk, Wang, Lei, Song, Hongbing, Pei, Tao, Jia, Zhongwei, Zeng, Daniel Dajun, 2020. Incorporating human movement data to improve epidemiological estimates for 2019-nCoV. medRxiv.

Fraser, Christophe, Riley, Steven, Anderson, Roy M., Ferguson, Neil M., 2004. Factors that make an infectious disease outbreak controllable. Proceedings of the National Academy of Sciences 101 (16), 6146–6151.

Wallinga, Jacco, Teunis, Peter, 2004. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. American Journal of epidemiology 160 (6), 509–516.

Wang, Yali, Wang, Xuan, Liu, Xiaobo, Ren, Ruiqi, Zhou, Lei, Li, Chao, Tu, Wenxiao, Ni, Daxin, Li, Qun, Feng, Zijian, et al., 2019. Epidemiology of imported infectious diseases, china, 2005-2016. Emerging infectious diseases 25 (1), 33–41.

Priest, Patricia C., Jennings, Lance C., Duncan, Alasdair R., Brunton, Cheryl R., Baker, Michael G., 2015. Effectiveness of border screening for detecting influenza in arriving airline travelers. American journal of public health 105 (S4), S607–S613.