Lessons drawn from China and South Korea for managing COVID-19 epidemic: insights from a comparative modeling study

Biao Tang 1,2, Fan Xia 1,4, Nicola Luigi Bragazzi 2, Zachary McCarthy 2,3, Xia Wang 3, Sha He 3, Xiaodan Sun 1,4, Sanyi Tang 1*, Yanni Xiao 1,4* & Jianhong Wu 1,2,5,6*

1 The Interdisciplinary Research Center for Mathematics and Life Sciences, Xi'an Jiaotong University, Xi'an 710049, People's Republic of China
2 Laboratory for Industrial and Applied Mathematics, Department of Mathematics and Statistics, York University, Toronto, Ontario, Canada, M3J 1P3
3 School of Mathematics and Information Science, Shaanxi Normal University, Xi'an, 710119, People's Republic of China
4 School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, People's Republic of China
5 Fields-CQAM Laboratory of Mathematics for Public Health, York University, Toronto, Ontario, Canada, M3J 1P3

The first three authors made the same contributions.
* Correspondence: sytang@snnu.edu.cn, yxiao@mail.xjtu.edu.cn, wujh@yorku.ca

(Submitted: 30 March 2020 – Published online: 1 April 2020)
Abstract: We conducted a comparative study of the COVID-19 epidemic in three different settings: mainland China, the Guangdong province of China and South Korea, by formulating two disease transmission dynamics models which incorporate epidemic characteristics and setting-specific interventions, and fitting the models to multi-source data to identify initial and effective reproduction numbers and evaluate effectiveness of interventions. We estimated the initial basic reproduction number for South Korea, the Guangdong province and mainland China as 2.6 (95% confidence interval (CI): (2.5, 2.7)), 3.0 (95%CI: (2.6, 3.3)) and 3.8 (95%CI: (3.5,4.2)), respectively, given a serial interval with mean of 5 days with standard deviation of 3 days. We found that the effective reproduction number for the Guangdong province and mainland China has fallen below the threshold 1 since February 8th and 18th respectively, while the effective reproduction number for South Korea remains high until March 2nd. Moreover our model-based analysis shows that the COVID-19 epidemics in South Korean is almost under control with the cumulative confirmed cases tending to be stable as of April 14th. Through sensitivity analysis, we show that a coherent and integrated approach with stringent public health interventions is the key to the success of containing the epidemic in China and especially its provinces outside its epicenter. In comparison, we find that the extremely high detection rate is the key factor determining the success in controlling the COVID-19 epidemics in South Korea. The experience of outbreak control in mainland China and South Korea should be a guiding reference for the rest of the world.

Keywords: COVID-19 epidemic; multi-source data; mathematical model; mainland China and South Korea; comparative study
Key points:

- We conducted a comparative study of the COVID-19 epidemic in three different settings: mainland China, the Guangdong province of China and South Korea.
- We estimated the initial basic reproduction number for South Korea, the Guangdong province and mainland China as 2.6 (95% confidence interval (CI): (2.5, 2.7)), 3.0 (95%CI: (2.6, 3.3)) and 3.8 (95%CI: (3.5,4.2)).
- A coherent and integrated approach with stringent public health interventions is the key to the success of containing the epidemic in China and especially its provinces outside its epicenter.
- The extremely high detection rate of infected individuals is the key to the success in controlling the epidemics in South Korea.
Introduction

Coronavirus, an enveloped virus characterized by a single-stranded, positive-sense RNA, causes generally mild infections but occasionally lethal communicable disorders leading to SARS, MERS\textsuperscript{1} and the current COVID-19 outbreak\textsuperscript{2,3} that has gradually spread out from the epicenter Wuhan/China and affected 103 countries/territories and international conveyances including the cruise ship Diamond Princess harbored in Yokohama/Japan as of March 7\textsuperscript{th} 2020.

In the absence of effective treatments and vaccines, an early adoption of stringent public health measures is crucial in mitigating the scale and burden of an outbreak. Unprecedented restrictive measures, including travel restrictions, contact tracing, quarantine and lock-down of entire towns/cities adopted by the Chinese authorities have resulted in a significant reduction of the effective reproductive number of COVID-19\textsuperscript{4,5}. However, these public health interventions may not be considered and/or implemented as effectively in other settings and contexts. Decision-making and implementations may require adaptations and modifications to take into account setting-specific characteristics in terms of community features, local epidemiology and risk assessment, social habits, juridical provisions, organizational coordination, and availability of economic-financial resources. For instance, particularly restrictive measures may not be effective in certain countries\textsuperscript{6}.

Several public health interventions can be implemented to counteract the threat posed by an emerging outbreak\textsuperscript{7} with pandemic potential. These interventions can be basically classified into two major categories: the measures in the first category are aimed at protecting the borders and include interventions such as travel restrictions and border entry screening, whereas the measures in the second category have the objective of locally controlling the spreading of the virus and include enhanced epidemiological surveys and surveillance, contact tracing, school closure and other interventions that favor a reduction in number of social contacts. The effectiveness of intervention measures is variable and, for some, is still under debate. Regarding, for example, extensive travel restrictions, a recent systematic review has shown that this intervention may contribute to delaying but not preventing the transmission and diffusion of a viral outbreak. As such, it is not recommended for implementation, if not within a broader package of public health measures aimed at rapidly containing the outbreak\textsuperscript{8}. A similar conclusion can be reached for border entry screening, considered as ineffective or poorly effective \textit{per se}, and therefore needs to be combined and provided together with other strategies\textsuperscript{9}. School closure appears to be potentially effective in containing/reducing viral outbreaks, although further research is warranted to identify the best strategy in terms of timing and length of closure\textsuperscript{10}. The measure of quarantine is also particularly controversial, since it raises ethical dilemmas, and political and social concerns\textsuperscript{11,12} and quantification of its real impact\textsuperscript{11} is difficult due to a high uncertainty in its efficacy. However,
in the absence of effective medical interventions, these measures must be implemented and the success of these measures, despite their disruptive impact on social-economic activities, depends heavily on how these measures are adapted to the specific scenario, in terms not only of clinical and epidemiological variables but also of social aspects, including social habits, juridical provisions, and economic-financial resources.

How differentiation and combination of these interventions within a coherent and systematic package of public health measures contributes to different outbreak outcomes is an urgent global health issue that must be addressed in order to ensure that lessons from countries that have early experienced COVID-19 outbreak can be learnt by other countries in their preparedness and management of a likely pandemic.

In South Korea, the first COVID-19 case (a 36 years old Chinese woman, with a recent travel history to Wuhan) was reported on January 8th 2020. A severe cluster of cases emerged in the city of Daegu, where on February 23rd 2020 a 61 years old woman spread the virus to hundreds of worshippers at Shincheonji Church of Jesus. On March 5th 2020, a further cluster of cases occurred at a nursing home in Gyeongsan, which has been declared “special care zone” in an effort to contain the viral outbreak. As of March 8th 2020, South Korea has reported 7,313 cases, with 130 total recovered cases and 50 deaths, with no sign that the epidemic is slowing down.

In comparison, intensive social contacts and massive mobility associated with the Chinese Traditional Spring Festival, combined with an initial delay in responding to the outbreak, resulted in an exponential growth of infections in the (then) epicenter (Wuhan) and large case importations to other Chinese cities. On January 23rd 2020, the Chinese government decisively implemented a systematic package of measures in the epicenter, including the lock-down/quarantine of Wuhan city and other cities/towns of the Hubei province, intense contact tracing and isolation. This led to rapid and effective mitigation of the COVID-19 epidemic. Case importation before the January 23rd lock-down also resulted in outbreaks in all Chinese provinces, but the systematic package of interventions implemented across the country led to effective containment. On March 8th 2020, the newly confirmed cases in the entire country reduced to 40.

Particularly, Guangdong, the province with the largest population in China at present, with GDP ranked first since 1989 and with the level of middle and upper income countries and middle developed countries, reported the first confirmed COVID-19 case on January 19th. On January 23rd, the government of Guangdong province announced the first-level response to major public health emergencies for controlling the spread of COVID-19. As of March 8th, there are totally 1325 confirmed cases and 8 deaths in Guangdong province, and no new case is reported. In contrast to South Korea, there was a relatively large ratio of imported cases in Guangdong province, particularly, in Shenzhen (a mega city in Guangdong province) more than 70 percent of the
confirmed cases are imported\textsuperscript{13}.

What differences in the intervention design and implementation between China and South Korea, and between Korea and Guangdong have contributed to the different epidemic curves, and what projection can we make about the epidemic trend in South Korea under different scenarios if a portion or the entirety of the Chinese package of public health interventions is applied to the South Korea? These are the questions we aim to answer by developing mathematical models tailored to different settings, including the entire country of China and its Guangdong province and South Korea.

**Methods**

**Data**

We obtained data of the confirmed COVID-19 cases, cumulative number of quarantined individuals, cumulative death cases in mainland China from the “National Health Commission” of the People’s Republic of China\textsuperscript{14}. Data information includes the newly reported cases, the cumulative number of reported confirmed cases, the cumulative number of cured cases and the number of death cases, as shown in Figure 1(A-C). In addition, we obtained the data of the cumulative confirmed cases, cumulative cured cases and daily cases under medical observation for the Guangdong province (Figure 1(E)) of China. We also obtained the data of cumulative confirmed cases and cumulative tested cases for South Korea from the Korea Centers for Disease Control and Prevention (KCDC)\textsuperscript{15,16}, as shown in Figure 1(D), (F). The data were released and analyzed anonymously. Note that the first confirmed case was reported on January 23\textsuperscript{rd} 2020 for South Korea, and also on January 23\textsuperscript{rd} 2020 mainland China started the lock-down of Wuhan city, the epicenter, and implemented other interventions. Note that the data for reported cases, either confirmed or quarantined, or under medical observation or tested, was used in China or South Korea since January 23\textsuperscript{rd} 2020.

**The model**

Our baseline model is the classical deterministic susceptible-exposed-infectious-removed (SEIR) epidemic model refined by incorporating contact tracing-quarantine-test-isolation strategies (Figure 2). We stratify the population into susceptible (\(S\)), exposed (\(E\)), symptomatic/asymptomatic infected (\(I/A\)), hospitalized (\(H\)) and recovered (\(R\)) compartments, and we further stratify the population to include quarantined susceptible (\(S_q\)), and quarantined suspected individuals (\(T\)). These stratifications were used in our previous studies\textsuperscript{4,5,17} and
agreement of model predictions with real data provides a validation of the model structure reflecting the interventions implemented in Wuhan and in mainland China. Here, we add an additional quarantined suspected compartment, which consists of exposed infectious individuals resulting from contact tracing and individuals with common fever. These individuals with common fever but quarantined as COVID-19 suspected contributed to the difficulty of implementing an effective quarantine process due to the size of this compartment. In what follows, exposure, transmission and infection compartments are always used for modeling the COVID-19.

In our model formulation, the transmission probability is denoted by $\beta$ and the contact rate is denoted by $c$. By enforcing contact tracing, a proportion, $q$, of individuals exposed is quarantined, and can either move to the compartment $B$ or $S_q$ with rate of $\beta c q$ (or $(1-\beta)cq$), depending on whether they are effectively infected or not\(^{18,19}\), while the other proportion, $1-q$, consists of individuals exposed to the virus who are missed from contact tracing and move to the exposed compartment $E$ at rate of $\beta c(1-q)$ once effectively infected or stay in compartment $S$ otherwise. Note that contact tracing is not triggered by asymptomatic infected individuals, who can infect susceptible individuals. We use constant $m$ to denote the transition rate from the susceptible compartment to the (COVID-19) suspected compartment due to fever and/or illness-like symptoms. The suspected individuals leave this compartment at a rate $b$, with a proportion, $f$, if confirmed to be infected with COVID-19, going to the hospitalized compartment, whilst the other proportion, $1-f$, is ruled out for COVID-19 infection and goes back to the susceptible compartment. Then, for mainland China, the model equations are as follows.

\[
\begin{align*}
S' &= -\left(\beta(t)c + \epsilon(t)q(t)(1-\beta)S\right) + \frac{\theta \beta c(t) SA}{N} - mS + \lambda S_q + b(1-f)T, \\
E' &= \frac{\beta c(t)(1-q(t))S}{N} + \frac{\theta \beta c(t) SA}{N} - aE, \\
I' &= \sigma E - (S(t) + \alpha)I, \\
A' &= \sigma(1-\rho)E - \gamma A, \\
T' &= \frac{\beta c(t)H(t)S}{N} + mS - bT, \\
S_q' &= \frac{(1-\beta)c(t)q(t)S}{N} - \lambda S_q, \\
H' &= \delta(t)I + b(T) - (\alpha + \gamma_H(t))H, \\
R' &= \gamma_A A + \gamma_H(t)H.
\end{align*}
\]

The effective reproduction number is given by

\[
R_t = \frac{\rho \beta c(t)(1-q(t))}{\delta(t)(1+a)} + \frac{\theta(1-\rho)\beta c(t)}{\gamma A}.
\]

In epidemiology, the basic reproduction number ($R_0$), used to measure the transmission potential of a disease, is the average number of secondary infections produced by a typical case of an infection in a population where everyone is susceptible. Once the intervention implemented results effective, the effective transmission potential is measured by the effective reproductive number $R_t$ that can be time-dependent as intervention measures vary over time. If $R_t$ falls below 1 and remains below 1, there will be a decline in the number of cases.
The prevention and control interventions were gradually improved in mainland China, and there are several key time points when mitigation measures were gradually strengthened: 1) On January 23rd Wuhan was locked down, and most parts of China shortly adopted a similar strategy; 2) On January 26th, the government announced to extend the Chinese Traditional New Year Festival holiday so self-isolation/protection was maximized; 3) On February 7th the Chinese government created the partnership between each one of the 16 provinces to its sister-city in the epicenter, the Hubei province, to reinforce the health care workers and equipment in the sister-city in Hubei; 4) On February 12th the Hubei province started to include the clinically diagnosed cases into the confirmed cases to enhance its quarantine/isolation measure; 5) On February 14th, Wuhan refined its management protocol of residential quarters; 6) On February 16th, the National Health Commission of the People’s Republic of China revised its New Coronavirus Pneumonia Prevention and Control Plan to further clarify and enhance the public health interventions in four key areas: Quarantine high-risk individuals as much as possible, Test suspected individuals as much as possible; Treat patients as best as possible; and Receive and cure all the patients.

We describe this improvement by using the time-dependent contact rate $c(t)$, quarantined rate $q(t)$, and detection rate $\delta(t)$, as follows

$$c(t) = (c_0 - c_b)e^{-r_1 t} + c_b, \quad q(t) = (q_0 - q_m)e^{-r_2 t} + q_m, \quad \frac{1}{\delta(t)} = \left(\frac{1}{\delta_{t0}} - \frac{1}{\delta_{tf}}\right)e^{-r_3 t} + \frac{1}{\delta_{tf}} \quad (C3)$$

where $c_0$ denotes the initial contact rate (on January 23rd 2020) with $c(0) = c_0$, $c_b$ denotes the minimal contact rate with the intervention being implemented. Constant $q_0$ is the initial quarantined rate of exposed individuals with $q(0) = q_0$, $q_m$ is the maximum quarantined rate with $q_m > q_0$. Similarly, constant $\delta_{t0}$ is the initial diagnose rate, $\delta_{tf}$ is the fastest diagnose rate with $\delta_{tf} > \delta_{t0}$.

For the Guangdong province, the recovery rate has been improved over the time, and this is reflected by the time-dependence of the recovery rate of the quarantined infected (isolated) population, given by

$$\gamma_H(t) = (\gamma_{H0} - \gamma_{Hf})e^{-r_4 t} + \gamma_{Hf} \quad (C4)$$

with $\gamma_{H0}$ being the initial recovery rate and $\gamma_{Hf}$ the fastest recovery rate, and $r_4$ the corresponding increasing rate of the recovery rate.

For South Korea, a similar contact tracing strategy has been implemented: all traced individuals are quarantined and are tested. Once tested, the suspected individuals leave the compartment at a rate $b$, with a proportion, $f$, if confirmed to be COVID-19 infected, going to the hospitalized compartment, whilst the remaining proportion, $1-f$, is recommended to remain quarantined and hence goes to the compartment $S_q^{18,19}$. The transmission dynamics subject to this intervention practice is given by
\[\begin{align*}
S' &= -\frac{(βc(t)(1-q)+c(t)q)SI}{N} - \frac{θβc(t)SA}{N} - mS + λS_q, \\
E' &= \frac{βc(t)(1-q)SI}{N} + \frac{θβc(t)SA}{N} - σE, \\
I' &= ρσE - (δ + α)I, \\
A' &= (1 - ρ)σE - γA, \\
T' &= \frac{c(t)SI}{N} + mS - bT, \\
S_q' &= b(1 - f)T - λS_q, \\
H' &= δI + bT - (α + γ)H, \\
R' &= γA + γH.
\end{align*}\]

(K1)

where \(c(t)\) has the same formula in equation (C3) while the other parameters are positive constants. The effective reproduction number of the model (K1) is given by

\[R_{tk} = \frac{ρβc(t)(1-q)}{δt+α} + \frac{θ(1-ρ)βc}{γA}\]

(K2)

**Model-free estimation for \(R_0\) and \(R_t\)**

We employed the method developed by White and Pagano\(^{20}\) to estimate the basic reproduction number \(R_0\). Let \(\{N_t\}_{t=0}^{T}\) be the number of newly reported cases on day 0,\,...,\,T. Assume \(N_t\) follows the Poisson distribution with a mean of \(E[N_t] = R_0 \sum_{j=1}^{t} p_j N_{t-j}\), where \(p_j(j = 1,2,\,...)\) is the probability, giving the serial interval distribution. Assume also that \(N_i\) and \(N_j\) are independent as long as \(i \neq j\). Given the distribution of the serial interval and the observed data on \(\{N_t\}_{t=0}^{T}\), the basic reproduction number \(R_0\) can be estimated by the maximum likelihood estimates approach.

With the same notations and assumptions described above, we can also estimate the effective reproduction number \(R_t\) following Cori\(^{21}\). Namely, using \(E[N_t] = R_t \sum_{j=1}^{t} p_j N_{t-j}\) within the Bayesian framework, we can obtain an analytical expression of the posterior distribution of \(R_t\) by assuming a gamma prior distribution for \(R_t\). Then we can get the posterior means and confidence intervals of \(R_t\).

**Results**

**Model-free estimation for \(R_0\) and \(R_t\)**

By using the number of daily newly reported cases from January 10\(^{th}\) to January 23\(^{rd}\) 2020, we estimate \(R_0\) for mainland China, and using the newly reported cases from January 19\(^{th}\) to January 31\(^{st}\) 2020 we estimate \(R_0\) for the Guangdong province. Also, we estimate \(R_0\) for South Korea based on the number of daily newly reported cases from January 23\(^{rd}\) to March 2\(^{nd}\) 2020.

All the estimates are given in Table 2. In particular, given the serial interval with mean of 5 and
standard deviation of 3, $R_0$ for mainland China, the Guangdong Province and South Korea is
estimated to be 3.8 (95%CI: (3.5, 4.2)), 3.0 (95%CI: (2.6, 3.3)) and 2.6 (95%CI: (2.5, 2.7)),
respectively. In particular, the initial COVID-19 reproduction rate in South Korea was smaller
than that in the Guangdong Province.

In order to investigate the variation of $R_0$ with respect to the serial interval, we carry out a
sensitivity analysis by changing the mean of the serial interval from 4 to 8 days, the standard
deviation (Std) from 3 to 5. The sensitivity analysis is reported in Table 2, and we notice that $R_0$
increases when the mean of the serial interval increases and we remark that serial interval
examined by recent studies is shorter than that earlier estimation$^{22-24}$. It also follows from Table 2
that increasing Std of the serial interval only slightly decreases the estimated $R_0$ for a given mean
of the serial interval.

We also estimate the effective reproduction numbers $R_t$ for the considered regions, using the
number of daily newly reported cases from the date the first case was reported until March 2nd
2020 (Figure 3). It shows that the effective reproduction number $R_t$ in mainland China and in its
Guangdong province has fallen below the threshold 1 since February 18th and February 7th, while
the effective reproduction number of South Korea remains very high, indicating that there is still
room for improving the interventions in South Korea.

**Model-based prediction and interventions efficacy evaluation**

By simultaneously fitting the model (C1) to the multiple source data on the cumulative number of
reported cases, deaths, quarantined and suspected cases in mainland China, we obtain estimations
for unknown parameters and initial conditions, listed in Table 1. The best fitting result is shown as
black curves in Figure 4 with the estimated baseline exponential decreasing rate ($r_1 = r_{10}$) in the
contact rate function. We then conduct a sensitivity analysis of the cumulative reported, death,
quarantined, suspected cases, and the infected (asymptomatic/symptomatic) individuals by
shrinking the exponential index $r_1$, representing the weakening of the control interventions
relevant to the contact rate. As shown in Figure 6, the numbers of cumulative reported, death,
quarantined, suspected cases, and the peak value of the infected all increase significantly. In
particular, with $r_1 = 0$ corresponding to no reduction of the contact rate from the initial period,
the cumulative confirmed cases increases by more than six times as of April 1st (~600,000 cases)
and the peak value of the infected will increase by more than 3 times, in comparison with the
actual situation under the strong control measures implemented by the Chinese government.

We also conduct a sensitivity analysis regarding the detection rate $\delta(t)$, by decreasing the value
of $r_3$. We obtain a similar conclusion that the cumulative confirmed cases would reach the
number of 350,000 cases as of April 1st with a constant detection rate (no improvement of detection), shown in Figure 6. As illustrated in Figure 6(F), we can observe that while decreasing the detection rate would not affect the decreasing trend of the effective reproduction number, it however postpones the time when the threshold value of 1 is reached. Therefore, the outcome in the mainland China, both in terms of the infections avoided and the timing when the outbreak begins to be under control, is the consequence of a systematic package of social distancing (self-isolation and self-protection), contacting tracing, and detection/diagnosis.

Similarly, by simultaneously fitting the proposed model (K1) to the cumulative number of reported and tested cases for South Korea, we obtain the estimations for the unknown parameters and initial conditions, listed in Table 1. The best fitting result is shown as black curves in Figure 7. Comparing the estimated values of the detection rate $\delta$ in China and South Korea, we find an important difference that the diagnose rate in South Korea (a constant value of 0.651) is much higher than that in China, being even higher than the maximal diagnose rate in China. This is in line with the real control interventions adopted by the Korean government, that is to say, the government has actively coordinated the production of testing kits of COVID-19, consequently, the testing kits have been always available, even during the peak period of the epidemic. Apparently, as shown in Figure 7(A-B), the COVID-19 epidemics in Korea will be almost under control as of April 14th with the confirmed cases tending to be stable. In order to test whether the extremely high detection rate is the key factor determining the success in controlling the COVID-19 in South Korea, we plotted the cumulative confirmed cases and the infectious population $I(t)$ by decreasing the detection rate by 70%, 50% and 30%, respectively, as shown in Figure 7(C-D). It follows from Figure 7(C) that the number of the cumulative reported cases can reach ~186,000 as of April 14th if the detection rate decreases by 70 percent (the red curve), which is around 17 times the predicted number of the best fitting result (the black curve). Furthermore, from Figure 7(D), we find that the peak value of the infectious population can increase significantly as well (around 30 times) if the detection rate is decreased by 70%, while the peak time is postponed a lot with a lower detection rate of the infectious individuals.

Further, by simultaneously fitting the model (C1) to the multiple source data on the cumulative number of reported cases, recovery and suspected cases of Guangdong province, we parameterize the model and obtain the estimations for the unknown parameters and initial conditions, listed in Table 1. The best fitting result is shown as green curves in Figure 8. Similarly, we consider the variation in the epidemic by means of the parameter $r_1$, representing the variation in the intensity of the control measures implemented. It follows from Figure 8(A-C) that increasing the parameter $r_1$ reduces the disease infections and the number of suspected individuals, while decreasing the parameter $r_1$ (to zero) slightly affects the disease infections, shown in green curves ($r_1 \neq 0$) and
red curves \( r_1 = 0 \) in Figure 8(A-C). This conversely implies that the prevention and control strategies in Guangdong province were relatively strong from the early stage of the outbreak. This can also be confirmed by the continuously declining trend of the effective reproduction number, as shown in Figure 3 and 8(D). Again, we performed the sensitivity analysis by decreasing the value of \( r_3 \) in the detection rate \( \delta_1(t) \) and obtained similar results, as shown in Figure 8.

**Discussion**

After some initial delay in the response to the COVID-19 outbreak, the Chinese government started implementing drastic public health measures on January 23rd 2020, including the lock-down of Wuhan city, contact tracing, quarantine, and isolation. This decisive and systematically implemented package of interventions has contributed to rapid mitigation, providing a window of opportunity for the preparedness, prevention and control of COVID-19 in other Chinese cities and around the world. As we showed in this modeling study, a systematic approach incorporating coherent and complementary public health measures rapidly implemented is the key for containing the viral outbreak.

In this study, we formulated a dynamic model and parameterized it by using multi-source data for the different settings within South Korea, China and Guangdong province. We predicted that the trend of cumulative reported cases would be very serious, reaching around 600 thousand as of April 1st 2020 given no control measures being implemented (Figure 5(A)). Hence a comparison with the current COVID-19 epidemic: the cumulative reported cases are around 80,651 as of March 6th, 2020, supporting the efficacy of the ongoing strengthening of prevention and control measures implemented in mainland China.

In Guangdong, the province with the largest population in China, most cases at the initial phase of the COVID-19 outbreak are imported, and immediately after the lockdown of Wuhan on January 23rd 2020, the province implemented a systematic approach towards prevention and control, with gradual enhancement, leading to the effective control of otherwise potentially catastrophic outcomes.

In comparison with South Korea, Guangdong has more inhabitants and a less developed economy. Also, from our model-free estimation, the basic reproduction number in South Korea is less than that computed for the Guangdong province. Therefore, the COVID-19 epidemic potential in South Korea was initially weaker than that in Guangdong. Our model-based analysis shows that the COVID-19 epidemics in South Korea is almost under control because of the high detection rate of the infectious individuals. Our simulation results indicate that the COVID-19 epidemic in South Korea will change from a slow to a quick increase in terms of the cumulative number of the
confirmed cases and the infected population if the detection rate is decreased, as shown in Figure 7(C-D). This indicates that the high detection rate is actually the key of the success in controlling the epidemics in South Korea. It is worth to mention that we assumed in our model that all the confirmed cases are isolated in the hospital, hence cannot infect other populations, which should be also a very important factor determining the success in controlling the COVID-19 epidemics. In other words, with the high detection rate, receiving and curing the patients as much as possible, instead of just asking the patients staying at home, is also important.

More in detail, a comparison of the parameter estimations of the Guangdong province and the entire country of China shows that

(1) the initial and maximum quarantine rates in Guangdong were much higher than those in the entire country China, while the initial and minimum contact rates were lower than those in the country, contributing to the observed better control effect in the province than the national average.

(2) the confirmation ratios of the Guangdong province and South Korea were much lower than the ratio of the entire country China, indicating the increased efficiency of contact tracing and testing in the Guangdong province and South Korea than that in the entire country of China.

Carrying out the sensitivity analysis of confirmed cases of China and Guangdong province with respect to the contact rate function, and further with respect to the detection rate of South Korea is in support of the effectiveness of the control measures implemented in other countries. Other countries or territories experiencing large-scale outbreaks of COVID-19 should capitalize on this lesson: convincing all the citizens to stay at home by self-isolation and becoming aware of the implications of self-protection is of crucial importance, but not enough. Improving the diagnosis rate of the infectious individuals and hospitalizing all the patients instead of staying at home have been fundamental for controlling the spreading of COVID-19 epidemic.

We note that future works may aim to address items of major public health concern such as: based on how a given regional epidemic situation is evolving, inform the decision-making for switching of intervention tactics and control measures. For instance, at which point does a country/territory make the decision to switch from tactics based on strict contact tracing to those based on travel restriction and mass quarantine? While contact tracing may be effective in controlling an outbreak with few initial cases, strong response efforts, and low pathogen transmissibility, this effectiveness may fade as cases accumulate. In other words, the effectiveness of control measures may be dependent on the time and epidemic state at which
they are implemented; therefore, switching tactics based on the local epidemic evolution may be optimal for regional and hence global control.

Through the parametrization and simulation of disease transmission models with intervention mechanisms and informed with multiple data sources, regional demographics, and regional intervention features, our work provides a foundational framework for the intervention evaluation and intervention scenario analysis. This study highlights the opportunity for evaluation of control measures and the learning from select regions to inform the future decision-making for the preparedness, real-time management as well as risk assessment of COVID-19.
Tables.

**Table 1: Parameter estimates for the COVID-19 epidemic in China, Guangdong Province of China and Korea.**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definitions</th>
<th>Estimated values</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c(t)$</td>
<td></td>
<td>Guangdong</td>
<td>China</td>
</tr>
<tr>
<td>$c_0$</td>
<td>Contact rate at the initial time</td>
<td>10.00</td>
<td>14.78</td>
</tr>
<tr>
<td>$c_b$</td>
<td>Minimum contact rate under the current control strategies</td>
<td>4.00</td>
<td>7.78</td>
</tr>
<tr>
<td>$r_1$</td>
<td>Exponential decreasing rate of contact rate</td>
<td>0.01</td>
<td>2</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Probability of transmission per contact</td>
<td>0.13</td>
<td>0.143</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Modification factor of the transmission probability of asymptotical infected individual</td>
<td>0.02</td>
<td>0.0437</td>
</tr>
<tr>
<td>$q(t)$</td>
<td></td>
<td>Guangdong</td>
<td>China</td>
</tr>
<tr>
<td>$q_0$</td>
<td>Quarantined rate of exposed individuals at the initial time</td>
<td>0.28</td>
<td>1.00 x 10^{-4}</td>
</tr>
<tr>
<td>$q_m$</td>
<td>Maximum quarantined rate of exposed individuals under the current control strategies</td>
<td>0.99</td>
<td>0.8</td>
</tr>
<tr>
<td>$r_2$</td>
<td>Exponential increasing rate of quarantined rate of exposed individuals</td>
<td>0.0361</td>
<td>0.15</td>
</tr>
<tr>
<td>$q$</td>
<td>Constant quarantined rate</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>$m$</td>
<td>Transition rate of susceptible individuals to the suspected class</td>
<td>$1 \times 10^{-5}$</td>
<td>$1 \times 10^{-6}$</td>
</tr>
<tr>
<td>$b$</td>
<td>Detection rate of the suspected class</td>
<td>0.0891</td>
<td>0.20</td>
</tr>
<tr>
<td>$f$</td>
<td>Confirmation ratio: Transition rate of exposed individuals in the suspected class to the quarantined infected class</td>
<td>0.01</td>
<td>0.50</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Ratio of symptomatic infection</td>
<td>0.5</td>
<td>0.4022</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Transition rate of exposed individuals to the infected class</td>
<td>1/5</td>
<td>1/5</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Rate at which the quarantined uninfected contacts were released into the wider community</td>
<td>1/14</td>
<td>1/14</td>
</tr>
<tr>
<td>$\delta_{1}(t)$</td>
<td>Initial transition rate of symptomatic infected individuals to the quarantined infected class</td>
<td>0.12</td>
<td>0.1326</td>
</tr>
<tr>
<td>$\delta_{1f}$</td>
<td>Fastest diagnose rate</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>$r_3$</td>
<td>Exponential decreasing rate of diagnose rate</td>
<td>0.2410</td>
<td>0.10</td>
</tr>
<tr>
<td>$\delta_{i}$</td>
<td>Constant transition rate of symptomatic infected individuals to the quarantined infected class</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>$\gamma_a$</td>
<td>Recovery rate of asymptotic infected individuals</td>
<td>0.1397</td>
<td>0.1397</td>
</tr>
<tr>
<td>$\gamma_0(t)$</td>
<td>Recovery rate of quarantined infected individuals at initial time</td>
<td>0.001</td>
<td>--</td>
</tr>
<tr>
<td>$\gamma_{1f}$</td>
<td>Fastest recovery rate of quarantined infected individual</td>
<td>0.2283</td>
<td>--</td>
</tr>
<tr>
<td>$r_4$</td>
<td>Exponential increasing rate of recovery rate of quarantined infected individuals</td>
<td>0.01</td>
<td>--</td>
</tr>
<tr>
<td>$\gamma_N$</td>
<td>Constant recovery rate of</td>
<td>--</td>
<td>0.2</td>
</tr>
</tbody>
</table>
quarantined infected individuals

<table>
<thead>
<tr>
<th>Initial values</th>
<th>Definitions</th>
<th>Estimated values</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Guangdong</td>
<td>China</td>
</tr>
<tr>
<td>$S(0)$</td>
<td>Initial susceptible population</td>
<td>$8.00 \times 10^5$</td>
<td>$3.00 \times 10^7$</td>
</tr>
<tr>
<td>$E(0)$</td>
<td>Initial exposed population</td>
<td>250</td>
<td>$3.00 \times 10^3$</td>
</tr>
<tr>
<td>$I(0)$</td>
<td>Initial infected population</td>
<td>378</td>
<td>$4.00 \times 10^3$</td>
</tr>
<tr>
<td>$A(0)$</td>
<td>Initial asymptotical population</td>
<td>100</td>
<td>$1.00 \times 10^3$</td>
</tr>
<tr>
<td>$T(0)$</td>
<td>Initial suspected population</td>
<td>258</td>
<td>1072</td>
</tr>
<tr>
<td>$S_q(0)$</td>
<td>Initial quarantined susceptible population</td>
<td>21(Est)</td>
<td>7347</td>
</tr>
<tr>
<td>$H(0)$</td>
<td>Initial quarantined infected population</td>
<td>53</td>
<td>771</td>
</tr>
<tr>
<td>$R(0)$</td>
<td>Initial recovered population</td>
<td>2</td>
<td>34</td>
</tr>
</tbody>
</table>

Note that, ‘Est’ means that the parameter values are estimated by fitting the models to the data when the source column indicates that they are not. ‘F’ means that the parameter values are fixed as the same as those estimated by fitting the proposed model to the multiple source data of China.
Table 2: Estimated basic reproduction number $R_0$ for mainland China, Guangdong province of China, Korea for various serial intervals.

<table>
<thead>
<tr>
<th>Std=3</th>
<th>E=4</th>
<th>E=5</th>
<th>E=6</th>
<th>E=7</th>
<th>E=8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mainland China</td>
<td>2.9(2.6,3.2)</td>
<td>3.8(3.5,4.2)</td>
<td>5.4(4.9,6.0)</td>
<td>7.8(7.1,8.6)</td>
<td>11.1(10.0,12.2)</td>
</tr>
<tr>
<td>Guangdong South Korea</td>
<td>2.4(2.1,2.7)</td>
<td>3.0(2.6,3.3)</td>
<td>3.8(3.4,4.3)</td>
<td>5.1(4.5,5.8)</td>
<td>6.9(6.1,7.8)</td>
</tr>
<tr>
<td>Guangdong South Korea</td>
<td>2.1(2.0,2.2)</td>
<td>2.6(2.5,2.7)</td>
<td>3.4(3.2,3.5)</td>
<td>4.5(4.3,4.7)</td>
<td>6.0(5.7,6.3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Std=4</th>
<th>E=4</th>
<th>E=5</th>
<th>E=6</th>
<th>E=7</th>
<th>E=8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mainland China</td>
<td>2.8(2.5,3.1)</td>
<td>3.4(3.1,3.7)</td>
<td>4.4(4.0,4.9)</td>
<td>6.0(5.4,6.6)</td>
<td>8.4(7.6,9.2)</td>
</tr>
<tr>
<td>Guangdong South Korea</td>
<td>2.4(2.1,2.7)</td>
<td>2.8(2.5,3.1)</td>
<td>3.4(3.0,3.9)</td>
<td>4.4(3.9,4.9)</td>
<td>5.8(5.1,6.5)</td>
</tr>
<tr>
<td>Guangdong South Korea</td>
<td>2.1(2.0,2.1)</td>
<td>2.4(2.3,2.5)</td>
<td>3.0(2.8,3.1)</td>
<td>3.8(3.7,4.0)</td>
<td>5.0(4.8,5.3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Std=5</th>
<th>E=4</th>
<th>E=5</th>
<th>E=6</th>
<th>E=7</th>
<th>E=8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mainland China</td>
<td>2.8(2.5,3.1)</td>
<td>3.2(2.9,3.5)</td>
<td>3.9(3.5,4.3)</td>
<td>4.9(4.5,5.6)</td>
<td>6.6(5.9,7.2)</td>
</tr>
<tr>
<td>Guangdong South Korea</td>
<td>2.4(2.1,2.7)</td>
<td>2.7(2.4,3.0)</td>
<td>3.2(2.8,3.6)</td>
<td>3.9(3.4,4.3)</td>
<td>4.9(4.3,5.5)</td>
</tr>
<tr>
<td>Guangdong South Korea</td>
<td>2.1(2.0,2.2)</td>
<td>2.3(2.2,2.4)</td>
<td>2.7(2.6,2.9)</td>
<td>3.4(3.2,3.5)</td>
<td>4.3(4.1,4.4)</td>
</tr>
</tbody>
</table>

E: mean of serial interval, Std: standard deviation of serial interval.

**Funding:** This research was funded by the National Natural Science Foundation of China (grant numbers: 11631012 (YX, ST), 61772017 (ST)), and by the Canada Research Chair Program (grant number: 230720 (JW)) and the Natural Sciences and Engineering Research Council of Canada (Grant number: 105588-2011 (JW)).

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


Figure 1. The datasets related to the COVID-19 epidemics, including newly reported cases, cumulative number of reported cases, cumulative number of cured cases, cumulative number of death cases, cumulative quarantined cases and cumulative suspected cases for mainland China (A-C), the Guangdong province of China (E), and South Korea (F).
Figure 2: The illustration of the compartmental models incorporating important interventions and features of reporting systems, for mainland China (model I) including its Province of Guangdong, and for South Korea (model II).
Figure 3: Estimated effective reproduction number $R_t$ over sliding weekly windows for the entire country of China, the Guangdong province of China, and Korea. The solid lines show the posterior means and the coloured zones show the 95% confidence intervals; the horizontal dashed line indicates $R_t=1$. 
Figure 4. Goodness of fit (black curves) and variations in cumulative number of reported cases (A), death cases (B), quarantined cases (C) and suspected cases (D) for mainland China. (E) Variation in number of infected (asymptomatic/symptomatic) individuals with contact rate function $c(t)$. (F) Here the contact rate function is changed by varying the exponential decreasing rate $r_1$, representing the variation in intensity of control measures. $r_{10}$ denotes the estimated baseline value of $r_1$. 
Figure 5. Goodness of fit (black curves) and variations in cumulative number of reported cases (A), cumulative number of death cases (B), quarantined cases (C) and suspected cases (D) for mainland China. (E) Variation in number of infected (asymptomatic/symptomatic) individuals with detection rate function $\delta_i(t)$. (F) Variation in the effective reproduction number of China. Here the detection rate function is changed by varying the exponential decreasing rate $r_3$, representing the variation in intensity of control measures. $r_{30}$ denotes the estimated baseline value of $r_3$. 
Figure 6. (A-B) Best fitting results and the impact of the randomness of the data set including the cumulative number of tested cases and the cumulative number of reported cases on the COVID-19 epidemic in Korea. The 95% confidence intervals have been given and the mean curve is marked as black. The red cycles denote the real data. (C-D) Sensitivity analysis. The impact of the diagnose rate $\delta_t$ on the cumulative confirmed cases and the infected populations in South Korea. Here, $\delta_{t0}$ denotes the estimated baseline value of the diagnose rate.
Figure 7. Goodness of fit (green curves) and variations in cumulative number of reported cases (A), recovery cases (B), and suspected cases (C) for Guangdong province. (D) Variation in the effective reproduction number with parameter $r_1$ in contact rate function $c(t)$. Here the contact rate function is changed by varying the exponential decreasing rate $r_1$, representing the variation in intensity of control measures. $r_{10}$ denotes the estimated baseline value for parameter $r_1$. 
Figure 8. Goodness of fit (green curves) and variations in cumulative number of reported cases (A), recovery cases (B), and suspected cases (C) for Guangdong province. (D) Variation in the effective reproduction number with parameter $r_3$ in detection rate function $\delta_i(t)$. Here the detection rate function is changed by varying the exponential decreasing rate $r_3$, representing the variation in intensity of control measures. $r^0_3$ denotes the estimated baseline value of $r_3$. 