


Análisis de datos visuales y predicción de simulación para COVID-19

**Baoquan Chen^a, Mingyi Shi^a, Xingyu Ni^a, Liangwang Ruan^a,
Hongda Jiang^a, Heyuan Yao^a, Mengdi Wang^a, Zhenhua Song^a,
Qiang Zhou^a, & Tong Ge^a**

Recibido: 2 de marzo de 2020 - Aceptado: 5 de marzo de 2020

Resumen: La epidemia de COVID-19 se ha convertido en una emergencia sanitaria mundial, tal y como declaró la WHO. China ha sido la más afectada por el brote de este virus, cuyo origen algunos expertos datan de finales de noviembre. No fue hasta el 23 de enero cuando el gobierno de Wuhan finalmente reconoció la gravedad de la epidemia y adaptó una medida drástica para frenar la propagación del virus, cerrando todos los transportes que la conectaban con el mundo exterior. En este estudio, buscamos responder algunas preguntas: ¿Cómo se propagó el virus desde el epicentro de la ciudad de Wuhan al resto del país? ¿Hasta qué punto las medidas como el cierre de la ciudad y la cuarentena de la comunidad ayudaron a controlar la situación? Y lo que es más importante, ¿podemos prever algún desarrollo futuro importante de este hecho si algunas de las condiciones hubieran cambiado? Al recopilar y visualizar los datos disponibles públicamente, mostramos primero las pautas y características del desarrollo de la epidemia; luego usamos un modelo matemático que refleja la dinámica de la transmisión de la enfermedad con el fin de evaluar la eficacia de algunas medidas de control de la epidemia y por último, lo que es más importante, para ofrecer algunos consejos sobre medidas preventivas.

Palabras clave: Simulación, COVID-19, Epidemia, Pandemia, Análisis de datos.

^a Center on Frontiers of Computing Studies, Peking University (China). Baoquan Chen  ORCID 0000-0002-8800-836X. Correspondence: Baoquan Chen, No.5 Courtyard of Jing Yuan, Peking University, No.5 Yiheyuan Road Haidian District, Beijing, China 100871. baoquan@pku.edu.cn

1. Escenario de la transmisión de la epidemia

El mapa de intensidad de la Fig. 1 visualiza la transmisión de la epidemia en su fase inicial, basado en el número de pacientes diagnosticados diariamente en toda China. Está claro que la epidemia se propagó desde Wuhan (Izquierda: el punto rojizo del centro) a sus alrededores, y más allá al resto del país, especialmente a las grandes áreas metropolitanas, como Beijing, Shanghai y Guangzhou (Derecha: puntos rojizos del Norte al Sur a lo largo de la costa), que se convirtieron en centros epidémicos secundarios en sus alrededores. Este patrón de transmisión refleja el gran movimiento de población antes del Año Nuevo Chino (25 de enero), cuando la gente suele regresar a casa para celebrar un Festival de Primavera de una semana de duración. El 23 de enero, la ciudad de Wuhan tomó y aplicó una medida extrema de cuarentena en toda la ciudad, cerrando todos los transportes que entraban y salían de la ciudad de Wuhan.

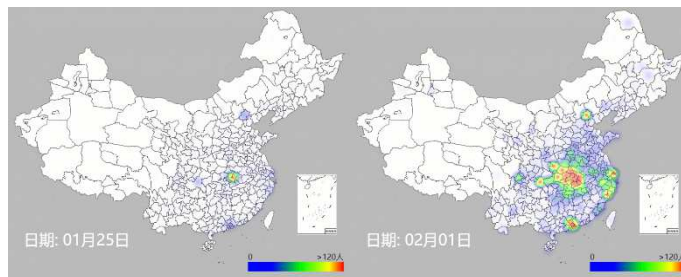


Figura 1. *Mapa de intensidad de la propagación de la epidemia en las provincias y ciudades de China (Der./Izq.: Ene 25/Feb 1)*

2. Visualización del patrón de propagación de la epidemia

Tomamos el período el 31 de enero, aproximadamente una semana después del cierre de la ciudad, como primera etapa del desarrollo epidémico, y posteriormente la segunda etapa. La duración de una semana es el tiempo promedio de incubación que una persona infectada tarda en mostrar síntomas. En la primera etapa, los nuevos casos de infección por el Coronavirus diagnosticados fuera de Wuhan se importaron principalmente de Wuhan a través de los movimientos de la población. Hemos recopilado datos de varias fuentes públicas (2019-nCoV Data Repository, 2020; Baidu Map, 2020) y visualizamos el volumen de población en el mapa.

A partir de la visualización (Fig. 2), podemos ver una fuerte correlación entre la población de infección diagnosticada y la migración desde Wuhan (centro-izquierda) en la primera etapa. Sin embargo, el desarrollo de la epidemia comienza a tomar diferentes ritmos en diferentes provincias. Esto refleja nuevos casos de transmisión de persona a persona, con diferentes tasas en diferentes provincias. A continuación, simulamos la dinámica epidémica para analizar de modo numérico con el objetivo de predecir los procedimientos de transmisión para proporcionar una evaluación de las diferentes medidas de control.

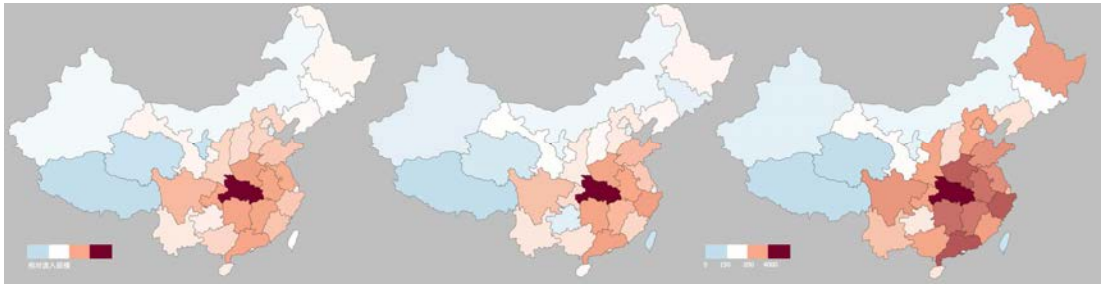


Figura 2. (Izquierda) la población fluye desde Wuhan al resto del país. Casos de infección diagnosticados el 31 de enero. Casos de infección diagnosticados el 9 de febrero.

3. Simulación de la dinámica de transmisión de epidemias

El 31 de enero, The Lancet publicó un artículo (Wu, Leung, & Leung, 2020) sobre la previsión y el pronóstico de la enfermedad COVID-19 (entonces llamada 2019-nCoV) que se estaba extendiendo por China. En este trabajo, los autores utilizaron el modelo clásico SEIR, que divide la población en cuatro categorías de individuos: susceptible (S), expuesto (E), infectado (I) y eliminado (R). El modelo supone que los individuos se transfieren entre categorías con una cierta probabilidad. Se estima que el número regenerativo R_0 es 2,68. A partir de este modelo, estimaron que las personas infectadas en Wuhan serían alrededor de 75.815 para el 25 de enero, y el número diario de infección sería de cientos de miles. Estimaron que otras varias ciudades altamente pobladas (Pekín, Shanghai, Chongqing, Guangzhou) presentaban una tendencia idéntica, excepto un par de días en las fechas del pico. El artículo básicamente descartó el efecto de la cuarentena en la ciudad de Wuhan y otras medidas de control agresivas en todo el país. El pronóstico describía un grave escenario tanto en China como en el mundo, resultaba útil para alertar al público, pero parecía estar bastante alejado de la realidad incluso en el momento de su publicación.

Uno de los principales problemas de la simulación mencionada es que no tiene en cuenta las agresivas medidas de control que se aplicaron en China,

donde: 1) se establecen procedimientos de prueba cada vez más eficaces para diagnosticar oportunamente a las personas infectadas, que inmediatamente son puestas en cuarentena, si es que no lo estaban ya; 2) las personas que tuvieron contacto con las personas infectadas también eran puestas en cuarentena; y 3) asimismo se pone en cuarentena a otro grupo de personas de las que se sospecha que pertenecen a las dos categorías mencionadas. Los individuos de las dos últimas categorías pueden ser dados de alta o trasladados a la primera categoría después de una posterior observación y prueba. Además de las mencionadas, existían más medidas para reducir la movilidad dentro de la ciudad.

3.1. Nuestro modelo

Tenemos como objetivo emplear un nuevo modelo de dinámica de la epidemia con la finalidad de incluir las medidas anteriormente mencionadas que se llevaron a cabo en China. Para ello, empleamos el modelo C-SEIR (Zhang, et al. 2005), que añade dos nuevos grupos de cuarentena en el SEIR (Fig. 3.1): el grupo de infección sospechosa en cuarentena (P) y el grupo de infección diagnosticada en cuarentena (Q). Los individuos de los grupos P y Q no tienen la capacidad de transmitir el virus.

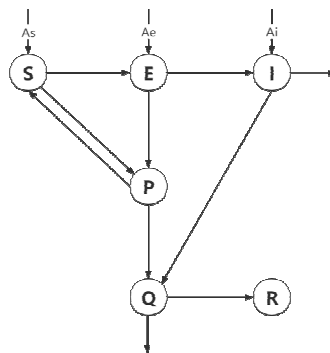


Figura 3.1. Modelo C-SEIR

La tasa de infección del virus $\beta(t)$ se define como el número diario de personas recién infectadas ΔE dividido por el número de pacientes no tratados I. Teniendo en cuenta que el nuevo virus también puede tener cierta infectividad durante el período de incubación, puede escribirse como:

$$\beta(t) = \frac{\Delta E}{I + kE}$$

Con k como 0,1, un parámetro empírico que significa que la capacidad infecciosa durante el período de incubación es 0,1 veces mayor que cuando se expresan los síntomas. Para determinar la forma específica de $\beta(t)$, primero calculamos el valor estimado de $\beta(t)$ utilizando el número diario de pacientes confirmados y el periodo de incubación estimado publicado por la Comisión Médica y de Salud de China para inferir el número diario de períodos de incubación y el número de infecciones, para después ajustar los datos con ciertas funciones elegidas. Con el fin de calcular la tasa infecciosa, es necesario calcular el número de individuos expuestos (E) y el número de pacientes infectados (I) cada día. Dado que únicamente se dispone de datos de pacientes recién diagnosticados, necesitamos estimar E e I en base a ellos. Al igual que en el artículo (Zhang, et al., 2005), suponemos que el tiempo entre dos generaciones de infección, y transcurrido entre la infección y el tratamiento del 2019-nCoV, son similares al SARS, que son 9 y 3 días respectivamente. Es decir, podemos asumir de manera aproximada que un paciente expuesto en el día t estará infectado en el día $t+6$, y será tratado en el día $t+9$. Por lo tanto, se puede estimar que el número total de pacientes admitidos durante $t \sim t + 9$ es igual al número de individuos expuestos en el día t , y el número total de pacientes admitidos durante $t \sim t + 3$ es igual al número de pacientes infectados en el día t . A partir de esto, se puede determinar E e I de cada día, y luego $\beta(t)$.

Durante la propagación de la mayoría de las enfermedades infecciosas, la tasa de infección $\beta(t)$ disminuye exponencialmente con el tiempo. Por este motivo usamos una función exponencial para ajustar $\beta(t)$ cada día. Tomando Beijing como ejemplo, el resultado ajustado se muestra en la figura 3.2. Al Aplicar un cálculo similar a los datos de otras provincias y ciudades hemos obtenido curvas de tasas de infección muy similares. Esto significa que las tasas de infección del virus han disminuido muy rápidamente, lo que demuestra la eficacia de las medidas de control.

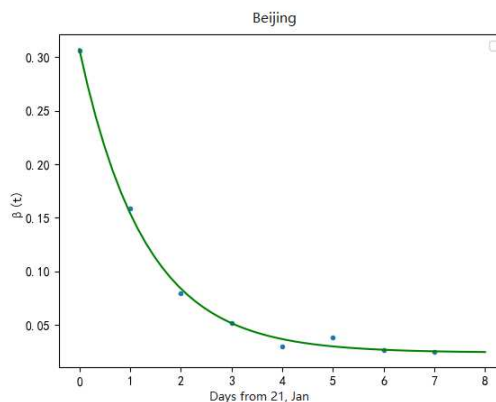


Figura 3.2. Curva ajustada del cambio de la tasa de infección en Beijing

3.2. Predicción de la simulación

Determinamos los parámetros del modelo y ejecutamos la simulación con el objetivo de ajustar el resultado en relación con los datos de diagnóstico confirmados diariamente y que fueron informados por la provincia de Hubei, cuya capital es Wuhan. A través de esto se puede obtener una predicción sobre el desarrollo futuro de la epidemia, que se muestra en la figura 3.3.

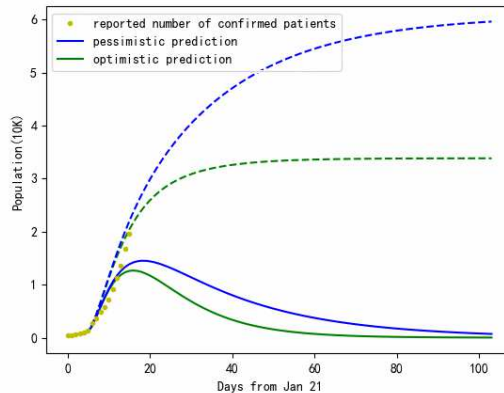


Figura 3.3. Predicción de infecciones (las curvas verdes y azules muestran los resultados de dos parámetros)

En este caso generamos dos curvas de predicción diferentes (azul / verde), donde la línea sólida es el número de pacientes confirmados diariamente y la línea discontinua el número acumulado de pacientes confirmados. En cualquier caso, la curva coincide con el número real de personas diagnosticadas (punto beige) en la etapa inicial. En la figura se puede ver que una ligera desviación en este ajuste de la curva en la etapa inicial puede llevar a una desviación significativa en el futuro. Sin embargo, la predicción "pesimista" (curva azul) da una fecha tope alrededor del 10 de febrero, siendo el número de pacientes confirmados acumulativamente de 50.000 el 1 de marzo, lo que se acerca bastante a las cifras oficiales publicadas. En marcado contraste, la predicción del artículo de *The Lancet* (Cohen, 2020) indica la magnitud de la realidad.

3.3 Evaluación de las medidas de control

Hemos preferido realizar simulaciones variando algunos parámetros para evaluar la eficacia de las diferentes medidas de control. En primer lugar, una cuarentena adecuada es lo más importante, especialmente en la fase inicial del desarrollo de la epidemia. Se cree que el brote de COVID-19 comenzó ya a mediados de noviembre, pero la ciudad de Wuhan no dio la alarma hasta el 20 de enero, con lo que no se aplicó una medida de cuarentena

efectiva hasta el 23 de enero. Como puede verse en la simulación (Fig 3.4), el cambio de la fecha de cuarentena, ya sea antes o después de 2 días, podría dar lugar a casi el doble de disminución o aumento de las personas infectadas de forma acumulativa. Es importante señalar que un número demasiado grande de personas infectadas puede agotar o incluso paralizar los recursos médicos disponibles. Lamentablemente, este fue el caso de la ciudad de Wuhan, que fue rescatada cuando llegaron recursos médicos (personal y suministros médicos) del resto del país. La ciudad también aplicó una cuarentena extremadamente agresiva en toda ella el 23 de enero, y una reducción de la movilidad interurbana a casi cero, con el objetivo de detener cualquier transmisión de la enfermedad.

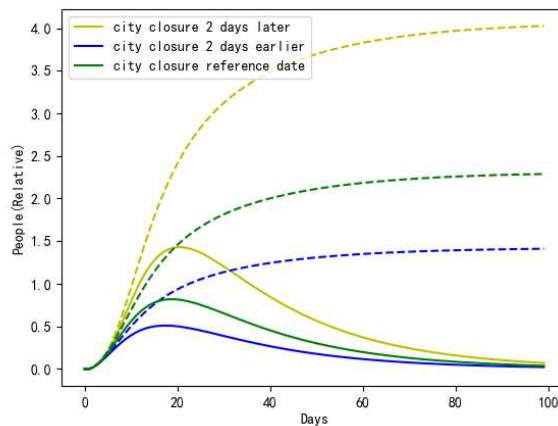


Figura 3.4. La simulación sugiere que el ejercicio de la cuarentena en la etapa inicial del brote de virus puede conducir a una doble disminución o aumento de las infecciones totales al cambiar la fecha de cuarentena sólo 2 días antes o después.

A partir de hoy (2 de marzo de 2020), la epidemia en China está muy controlada. Fuera de Wuhan se ha visto desde cero a un solo dígito diario confirmado de infección durante una semana más o menos. Esto se ha logrado mediante la mencionada medida de cuarentena agresiva en todo el país, tal como se describe en el informe de la WHO-China Joint Mission (WHO-China, 2020).

A continuación, examinamos el impacto de la flexibilización de las medidas de cuarentena en el desarrollo de la epidemia. En la figura 3.5 se muestran las predicciones sobre diferentes medidas de cuarentena: si la cuarentena se cancela completamente después de alcanzar el pico diario, la tasa de mitigación de la epidemia se reducirá en gran medida, lo que dará lugar a un posible segundo pico de infección (azul), mientras que una cuarentena parcialmente relajada sólo da lugar a un ligero aumento de la infección (beige), en comparación con la medida de cuarentena normal (verde), lo que

ilustra tanto la importancia de la cuarentena para evitar un segundo pico como la necesidad de ejercer una medida equilibrada para mantener la normalidad de la vida y el trabajo en la medida de lo posible.

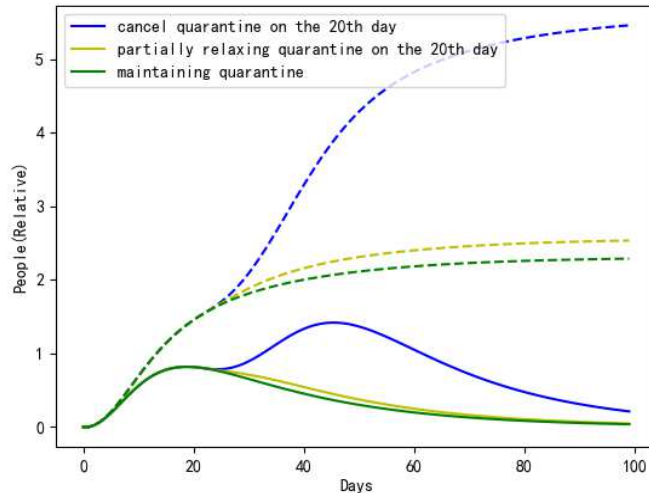


Figura 3.5. Predicción sobre diferentes medidas de cuarentena: la cancelación completa de la cuarentena (beige) da lugar a un segundo pico, mientras que una cuarentena parcialmente relajada sólo da lugar a un ligero aumento de las personas infectadas (azul), en comparación con las medidas de cuarentena normales.

4. Concluding Remarks with Tips Offered

Nuestra simulación basada en el modelo C-SEIR ha mostrado una efectividad significativa de las medidas de cuarentena, que ha sido aplicada a fondo en China. Si bien aplaudimos el éxito de estos esfuerzos, debemos señalar que esto tiene un precio demasiado alto. En retrospectiva, podría haberse evitado ejerciendo una cuarentena más apropiada (se podría decir menos estricta), pero mucho antes.

Dado que el COVID-19 parece estar extendiéndose a nivel mundial, hacemos algunas sugerencias basadas en nuestro análisis y observaciones hasta ahora:

- (1) Alert and educate the public about the virus and its prevention (not much different than preventing flu).
- (2) Vigilar de cerca el desarrollo de la epidemia diagnosticando rápidamente a las personas infectadas con síntomas característicos, ponerlas en cuarentena y tratarlas.
- (3) Alertar a aquellos que puedan haber tenido un contacto cercano con los individuos infectados y pedirles que se mantengan en una actividad social baja aunque no tengan síntomas, o que hagan

autocuarentena si los síntomas son leves, o que vayan al hospital para ser diagnosticados si los síntomas son los característicos. En la autocuarentena es mejor para permanecer hasta 2 semanas aislado.

La tercera medida refleja la estrategia de cuarentena única del modelo C-SEIR, y se ha demostrado que es eficaz tanto en la simulación del modelo como en su verificación en la realidad.

Un artículo reciente de expertos en salud de primera línea en China informó de una tasa de mortalidad del 1,4% (Guan, et al., 2020; Chen, et al. 2020). Pero fuera de la provincia de Hubei, la tasa de mortalidad parece ser mucho más baja que esto, hasta el 0,85%. Se puede especular que la letalidad del virus disminuye con las nuevas transmisiones. Las infecciones que se producen fuera de China son nuevas generaciones de la transmisión, por lo tanto, la tasa de mortalidad podría ser aún menor. Es posible que la transmisibilidad del virus también pueda reducirse con el tiempo. Sin embargo, el nuevo virus es definitivamente más severo que una gripe fuerte; por lo tanto, merece una especial precaución y contar con muchos recursos médico de antemano.

Referencias

- 2019-nCoV data repository (2020). *2019-nCoV data repository*. Retrieved from <https://github.com/BlankerL/DXY-2019-nCoV-Data>
- Baidu Map - migration (2020). *Baidu Map - migration*. Retrieved from <http://qianxi.baidu.com/>
- Chen, B., Shi, M., Ni, X., Ruan, L., Jiang, H., Yao, H., Wang, M., Song, Z., Zhou, Q., y Ge. T. (2020). Data Visualization Analysis and Simulation Prediction for COVID-19. *arXiv:2002.07096* [physics, q-bio], February 2020. Recuperado de <http://arxiv.org/abs/2002.07096>.
- Cohen, J. (2020). Scientists are racing to model the next moves of a coronavirus that's still hard to predict. *Science*, Feb. 7, 2020. Recuperado de <https://www.sciencemag.org/news/2020/02/scientists-are-racing-model-next-moves-coronavirus-thats-still-hard-predict>
- Guan, W. J., Ni, Z. Y., Hu, Y., Liang, W. H., Ou, C. Q., He, J. X., et al. (2020). Clinical Characteristics of Coronavirus Disease 2019 in China. *The New England Journal of Medicine*, February 28, 2020. doi: 10.1056/NEJMoa2002032; Recuperado de <https://www.nejm.org/doi/full/10.1056/NEJMoa2002032>
- WHO-China (2020). *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)*. Recuperado de <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>

- Wu, J. T., Leung, K., y 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*, In Press, January 2020. doi: 10.1016/S0140-6736(20)30260-9. Recuperado de <https://linkinghub.elsevier.com/retrieve/pii/S0140673620302609>.
- Zhang. J., Lou, J., Ma. Z., et al. (2005). A compartmental model for the analysis of SARS transmission patterns and outbreak control measures in China. *Applied Mathematics and Computation*, 162(2), 909-924.



© 2020 Chen, Shi, Ni, Ruan, Jiang, Yao, Wang, Song, Zhou, & Ge; licensee International Journal of Educational Excellence, Universidad Ana G. Méndez (UAGM). This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.


Visual Data Analysis and Simulation Prediction for COVID-19

**Baoquan Chen^a, Mingyi Shi^a, Xingyu Ni^a, Liangwang Ruan^a,
Hongda Jiang^a, Heyuan Yao^a, Mengdi Wang^a, Zhenhua Song^a,
Qiang Zhou^a, & Tong Ge^a**

Received: 2 March 2020 • Accepted: 5 March 2020

Abstract: The COVID-19 (formerly, 2019-nCoV) epidemic has become a global health emergency, as such, WHO declared PHEIC. China has taken the most hit since the outbreak of the virus, which could be dated as far back as late November by some experts. It was not until January 23rd that the Wuhan government finally recognized the severity of the epidemic and took a drastic measure to curtail the virus spread by closing down all transportation connecting the outside world. In this study, we seek to answer a few questions: How did the virus get spread from the epicenter Wuhan city to the rest of the country? To what extent did the measures, such as, city closure and community quarantine, help controlling the situation? More importantly, can we forecast any significant future development of the event had some of the conditions changed? By collecting and visualizing publicly available data, we first show patterns and characteristics of the epidemic development; we then employ a mathematical model of disease transmission dynamics to evaluate the effectiveness of some epidemic control measures, and more importantly, to offer a few tips on preventive measures.

Key Words: Simulation, COVID-19, Epidemic, Pandemic, Data Analysis

^a Center on Frontiers of Computing Studies, Peking University (China). Baoquan Chen  ORCID 0000-0002-8800-836X. Correspondence: Baoquan Chen, No.5 Courtyard of Jing Yuan, Peking University, No.5 Yiheyuan Road Haidian District, Beijing, China 100871. baoquan@pku.edu.cn

1. Overview of the Epidemic Transmission

The heat map in Fig 1 visualizes the epidemic transmission in its early stage, based on daily diagnosed patient numbers across China. It's clear that the epidemic spread from Wuhan (Left: the hotspot in the middle) to its surrounding areas, and further to the rest of the country, especially the large metropolitan areas, such as Beijing, Shanghai, and Guangzhou (Right: hotspots from the North to South along the coast), which became secondary epidemic centers to its vicinities. Such transmission pattern reflects the large population movement prior to the Chinese New Year (January 25), when people usually return home to celebrate a week-long Spring Festival. On January 23, the Wuhan city exercised an extreme whole-city quarantine measure by shutting down all transportations going in and out of the Wuhan city.

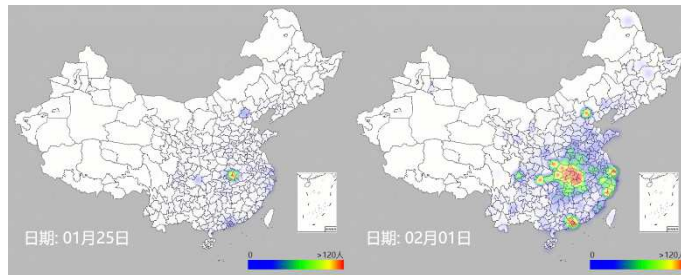


Figure 1. Heat map of epidemic spread in provinces and cities in China (L/R:Jan 25/Feb 1)

2. Visualization of Epidemic Spreading Pattern

We take the period January 31, approximately a week after the city closure, as the first stage of the epidemic development, after that, the second stage. The one-week duration is the averaged incubation time that it takes an infected person to show symptoms. At the first stage, the new Coronavirus infection cases diagnosed outside Wuhan are mainly imported from Wuhan through population movement. We collect data from several public sources (2019-nCoV data repository, 2020; Baidu Map, 2020) and visualize population volume on map.

From the visualization (Fig. 2), we can see strong correlation between the population of diagnosed infection and migration from Wuhan (Left-Middle) in the first stage, however, the epidemic development start to take different paces at different provinces. This reflects new cases of human-to-human transmission, with varying rates in different provinces. Next, we simulate epidemic dynamics to numerically analyze and predict the

transmission procedures, and provide evaluation on different control measures.



Figure 2. (Left) population flow from Wuhan to the rest of the country. (Middle) diagnosed infection cases in Jan. 31. (Right) diagnosed infection cases in Feb. 9.

3. Simulation of Epidemic Transmission Dynamics

On January 31, The Lancet published an article (Wu, Leung, & Leung, 2020) on nowcasting and forecasting of COVID-19 disease (then called 2019-nCoV) spreading in China. In this work, the authors used the classic SEIR model, which divides the population into four categories of individuals: susceptible (S), exposed (E), infected (I), and removed (R). The model assumes that individuals transfer between categories with a certain probability. The regenerative number R_0 is estimated to be 2.68. From this model, they estimated that the number of infected people in Wuhan had reached about 75,815 by January 25, and a daily infection number to be hundreds of thousands. They estimated several other populous cities (Beijing, Shanghai, Chongqing, Guangzhou) to have somewhat identical trend, except a couple of days shift on the peaking dates. The article basically dismissed the effect of the city-wide quarantine of Wuhan and other aggressive control measures around the country. The forecasting depicted a grave scenario both within China and globally, was helpful in alerting the public, but seemed to be quite off from the reality even at the time of its publishing.

One major problem of the above simulation is that it does not consider the aggressive control measures conducted around China, where (1) increasingly efficient testing procedures are put in place to timely diagnose infected individuals, who are then immediately quarantined, if not already; (2) individuals who had contact with the infected people are also quarantined; and (3) a further group of individuals who are suspected to be in the above two categories are also quarantined. The individuals in the latter two categories may be either released or moved to the first category after further observation and testing. In addition to these, there are more measures on reducing mobility within the city.

3.1. Our model

We aim to employ a new epidemic dynamics model to better consider the above measures in China. To this end, we employ the C-SEIR model (Zhang, et al. 2005), which has added two new quarantined groups in SEIR (Fig. 3.1): the quarantined suspected infection group (P), and the quarantined diagnosed infection group (Q). Individuals in groups P and Q do not have the ability to transmit the virus onward.

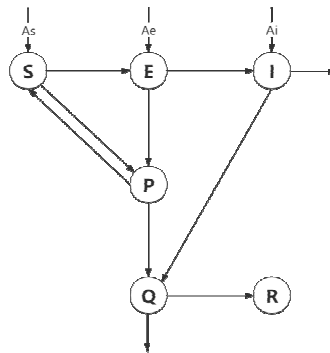


Figure 3.1. C-SEIR model

The virus infectious rate function $\beta(t)$ is defined as the daily number of newly infected people ΔE divided by the number of untreated patients I. Considering that the new virus may also have a certain infectivity during the incubation period, it can be written as:

$$\beta(t) = \frac{\Delta E}{I + kE}$$

With k as 0.1, an empirical parameter meaning that the infectious capacity during the incubation period is 0.1 times that when symptoms are expressed. In order to determine the specific form of $\beta(t)$, we first calculate the estimated value of $\beta(t)$ using the daily number of confirmed patients and the estimated incubation period published by the Health and Medical Commission of China to infer the daily number of incubation periods and the number of infections, and then fit the data with certain chosen functions.

In order to compute the infectious rate $\beta(t)$, it is necessary to calculate the number of exposed individuals (E) and the number of infected patients (I) for each day. Since only data of newly diagnosed patients are available, we need to estimate E and I based on it. Like in the article (Zhang, et al., 2005)

we assume that the time between two generations of infection, and the time between infection and treatment of 2019-nCoV, are similar to SARS, which are 9 days and 3 days, respectively. That is to say, we can roughly assume that a patient exposed at day t will be infected at day $t+6$, and be treated at day $t+9$. Thus, it can be estimated that the total number of patients admitted during $t \sim t + 9$ is equal to the number of exposed individuals on day t , and the total number of patients admitted during $t \sim t + 3$ is equal to the number of infected patients on day t . From this, we can estimate E and I of each day, then $\beta(t)$.

During the spread of most infectious diseases, the infectious rate $\beta(t)$ decreases exponentially with time. We then use an exponential function to fit $\beta(t)$ for each day. Taking Beijing as an example, the fitted result is shown in Figure 3.2. By applying similar calculation to the data from other provinces and cities, we have obtained largely similar infection rate curves. This means that the virus infection rates have dropped very quickly, a testimony to the effectiveness of the control measures.

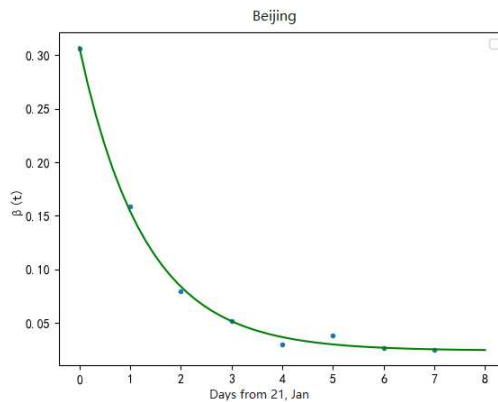


Figure 3.2. Fitted curve of infection rate change in Beijing

3.2. Simulation prediction

We estimate model parameters and run the simulation, aiming to fit the output with the daily confirmed diagnose data reported from Hubei Province, Wuhan being its capital city. Through this, we obtain prediction on the future epidemic development, shown in Figure 3.3.

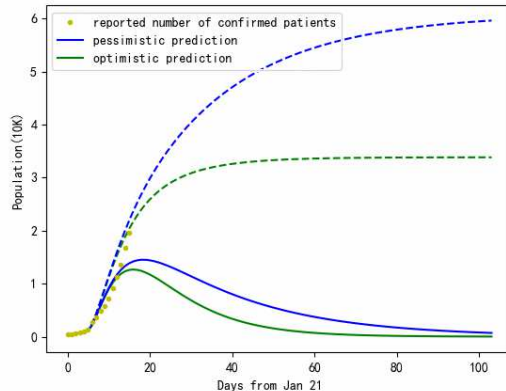


Figure 3.3. Infection prediction (green and blue curves show results of two parameters)

Here, we generate two different prediction curves (blue / green), where the solid line is the number of daily confirmed patients and the dashed line the cumulative number of confirmed patients. In either case, the curve fits with the actual number of people diagnosed (beige point) in the early stage. It can be seen from the figure that slight deviation on this curve fitting in the early stage may lead to significant departure going forward. Nevertheless, the ‘pessimistic’ prediction (blue curve) gives a peak date around Feb 10, and the cumulatively confirmed patients number to be 50K on March 1st, which is quite close to the official numbers publicized. In stark contrast, the prediction of the Lancet article (Cohen, 2020) is an order of magnitude that of the reality.

3.3 Evaluation of Control Measures

We like to conduct simulation by varying a few parameters to evaluate the effectiveness of different control measures.

First of all, proper quarantine is the most important, especially at the early stage of the epidemic development. It is believed that the outbreak of COVID-19 started as early as mid-November, but Wuhan city did not alert the public until Jan 20, and exercise effective quarantine measure until January 23rd. As can be seen from the simulation (Fig 3.4), shifting the quarantine date either earlier or later by 2 days could result in almost double amount of decrease or increase of the infected people cumulatively. It is important to note that too large a number of infected people may exhaust or even paralyze available medical resource. This unfortunately became the case for the Wuhan city, which was then rescued by pouring medical resources (medical personnel and supplies) from the rest of the country. The city also exercised an extremely aggressive whole-city quarantine on Jan 23, and a reduction of inter-city mobility to almost zero, with the aim of halting any disease transmission.

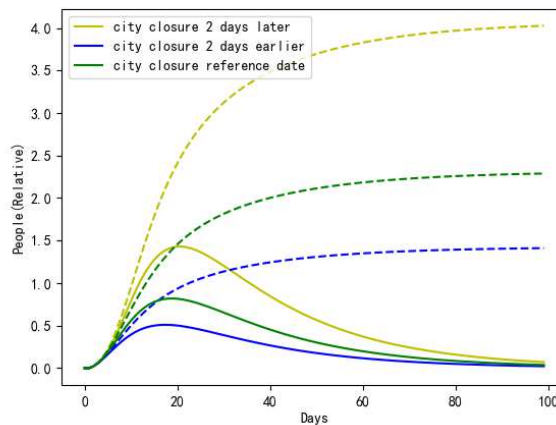


Figure 3.4. Simulation suggests that exercising quarantine at the early stage of the virus outbreak can lead to double decrease or increase of the total infections by shifting the quarantine date only 2 days earlier or later.

As of today (March 2, 2020), the epidemic in China is well under control. Outside Wuhan has seen zero to single digit daily confirmed infection for a week or so. This is achieved through the aforementioned aggressive quarantine measure across the nation, as described in the report by the WHO-China Joint Mission (WHO-China, 2020).

Next, we examine the impact of relaxing quarantine measures on epidemic development. Figure 3.5 shows predictions on different quarantine measures: if the quarantine is cancelled completely after reaching the daily peak, the rate of epidemic mitigation will be greatly reduced, resulting in a possible second peak of infection (blue), while a partially relaxed quarantine leads to only slight increase of infection (beige), comparing with the normal quarantine measure (green). This illustrates both the importance of quarantine to avoid a second strike, and the necessity of exercising a balanced measure to maintain life and work normalcy as much as possible.

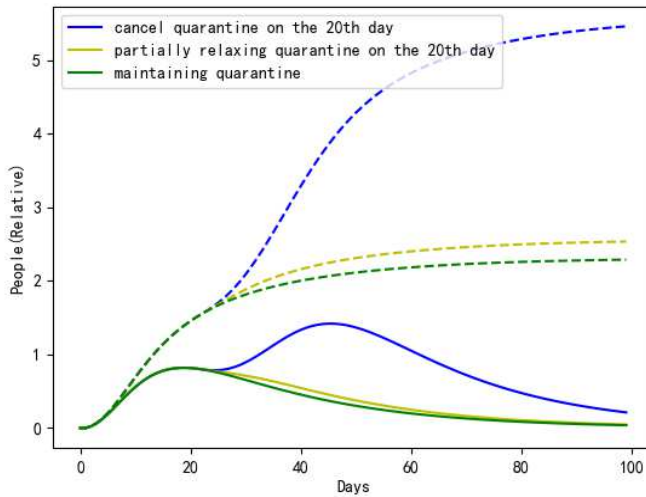


Figure 3.5. Prediction on different quarantine measures: cancelling quarantine completely (beige) results in a second peak, while a partially relaxed quarantine leads to only slight increase of the infected people (blue), comparing with the normal quarantine measures.

4. Concluding Remarks with Tips Offered

Our simulation based on the C-SEIR model has shown significant effectiveness of quarantine measures, which has been thoroughly exercised in China. While we applaud the success of these efforts, we must point out that this comes at a price much too high. In the hindsight, this could have been avoided by exercising a more proper (one could say less strict) quarantine, but much earlier.

Since the COVID-19 virus seems to be spreading globally, we make a few suggestions based on our analysis and observations so far:

- (1) Alert and educate the public about the virus and its prevention (not much different than preventing flu).
- (2) Closely monitor the epidemic development by quickly diagnosing infected people with signature symptoms, quarantine and treat them.
- (3) Alert those who may have had close contact with the infected individuals and ask them to stay low in social activity even if no symptoms, or exercise self-quarantine if light symptoms, or go to hospital for diagnosis if signature symptoms. The self-quarantine is best to stay for as long as 2 weeks.

The third measure reflects the unique quarantine strategy of the C-SEIR model, and has been proven effective both in the model simulation and the reality check.

A recent article by the frontline health experts in China reported a death rate of 1.4% (Guan, et al., 2020; Chen, et al. 2020). But outside Hubei province, the death rate seems to be much lower than this, down to 0.85%. A speculation could be that the lethality of the virus goes down with further transmissions. The virus infections going on outside China are further generations of the transmission, hence, the death rate could be even lower. It is speculated that the transmissibility of the virus may also reduce over time. Nevertheless, the COVID-19 virus is definitely more severe than a strong flu; therefore, it deserves extra precaution and well prepared medical resource in advance.

References

- 2019-nCoV data repository (2020). *2019-nCoV data repository*. Retrieved from <https://github.com/BlankerL/DXY-2019-nCoV-Data>
- Baidu Map - migration (2020). *Baidu Map - migration*. Retrieved from <http://qianxi.baidu.com/>
- Chen, B., Shi, M., Ni, X., Ruan, L., Jiang, H., Yao, H., Wang, M., Song, Z., Zhou, Q., & Ge, T. (2020). Data Visualization Analysis and Simulation Prediction for COVID-19. *arXiv:2002.07096* [physics, q-bio], February 2020. Retrieved from <http://arxiv.org/abs/2002.07096>.
- Cohen, J. (2020). Scientists are racing to model the next moves of a coronavirus that's still hard to predict. *Science*, Feb. 7, 2020. Retrieved from <https://www.sciencemag.org/news/2020/02/scientists-are-racing-model-next-moves-coronavirus-thats-still-hard-predict>
- Guan, W. J., Ni, Z. Y., Hu, Y., Liang, W. H., Ou, C. Q., He, J. X., et al. (2020). Clinical Characteristics of Coronavirus Disease 2019 in China. *The New England Journal of Medicine*, February 28, 2020. doi: 10.1056/NEJMoa2002032; Retrieved from <https://www.nejm.org/doi/full/10.1056/NEJMoa2002032>
- WHO-China (2020). *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)*. Retrieved from <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>
- Wu, J. T., Leung, K., & 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*, In Press, January 2020. doi: 10.1016/S0140-6736(20)30260-9. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S0140673620302609>.

Zhang. J., Lou, J., Ma. Z., et al. (2005). A compartmental model for the analysis of SARS transmission patterns and outbreak control measures in China. *Applied Mathematics and Computation*, 162(2), 909-924.



© 2020 Chen, Shi, Ni, Ruan, Jiang, Yao, Wang, Song, Zhou, & Ge; licensee International Journal of Educational Excellence, Universidad Ana G. Méndez (UAGM). This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.