

Coronavirus disease 2019 outbreak in Beijing's Xinfadi Market: a modeling study to inform future resurgence response

Xiaoli Wang^{1,2,3,a}, Xin Lin^{4,a}, Peng Yang^{1,2,a}, Zunyou Wu⁵, Gang Li¹, Jennifer M. McGoogan⁵, Zengtao Jiao⁶, Xinjun He⁶, Siqi Li⁷, Honghao Shi⁴, Jingyuan Wang^{4,*}, Shengjie Lai^{8,9,*}, Chun Huang^{1,*}, Quanyi Wang^{1,3,*}

Author affiliations

¹Beijing Research Center for Preventive Medicine, Beijing Center for Disease Prevention and Control, Beijing, China. ²School of Public Health, Capital Medical University, Beijing, China. ³Beijing Municipal Key Laboratory of Clinical Epidemiology, Beijing, China. ⁴School of Computer Science and Engineering, Beihang University, Beijing, China. ⁵Chinese Center for Disease Control and Prevention, Beijing, China. ⁶Yidu cloud AI lab, Yidu Cloud (Beijing) Technology Co., Ltd, Beijing, China. ⁷Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA. ⁸WorldPop, School of Geography and Environmental Science, University of Southampton, Southampton, UK. ⁹School of Public Health, Fudan University, Key Laboratory of Public Health Safety, Ministry of Education, Shanghai, China.

^aThese authors contributed equally to this work

***Corresponding authors**

Jingyuan Wang, PhD, School of Computer Science and Engineering, Beihang University, Beijing, China, jywang@buaa.edu.cn; Quanyi Wang, MD, Beijing Center for Disease Prevention and

Control, Beijing, China, bjcdexm@126.com; Chun Huang, MD, Beijing Center for Disease Prevention and Control, Beijing, China, huangch@bjcdc.org; Shengjie Lai, PhD, WorldPop, School of Geography and Environmental Science, University of Southampton, Southampton, UK, shengjie.lai@soton.ac.uk.

Word count

294 words abstract, 3,677 words main text, 22 references, 4 tables, 2 figures

Abstract

Background: A local coronavirus disease 2019 (COVID-19) case confirmed on June 11, 2020 triggered an outbreak in Beijing after 56 consecutive days without a newly confirmed case. Non-pharmaceutical interventions were used to contain the source in Xinfadi (XFD) market. To rapidly control the outbreak, both traditional and newly introduced NPIs including large-scale management of high-risk populations and expanded severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) PCR-based screening in the general population were conducted in Beijing.

Methods: A modified susceptible–exposed–infectious–recovered (SEIR) model was developed and used to evaluate a range of different scenarios from the public health perspective. Two outcomes were measured: magnitude of transmission (i.e., number of cases in the outbreak) and endpoint of transmission (i.e., date of containment). The outcomes of scenario evaluations were presented relative to the reality case (i.e., 368 cases in 34 days) with 95% credible intervals.

Results: Our results indicated that a 3 to 14-day delay in the identification of XFD as the infection source and initiation of NPIs would have caused a 3 to 28-fold increase in total case number and a 31 to 77-day delay in containment. A failure to implement the quarantine scheme employed in the XFD outbreak for defined key population would have caused a 5-fold greater number of cases and 73-day delay to containment. Similarly, failure to implement the quarantine plan executed in the XFD outbreak for close contacts would have caused 2-fold greater transmission and a 44-day delay to containment. Finally, failure to implement expanded nucleic acid screening in the general population would have yielded 1.6-fold greater transmission and a 32-day delay to containment.

Conclusions: The evidence provided by this study should inform responses to future outbreaks of COVID-19 and future infectious disease outbreak preparedness efforts in China and elsewhere.

Keywords: public health, nonpharmaceutical intervention, COVID-19, SARS-CoV-2, China

Background

Although it has been more than 12 months since the first confirmed case of novel coronavirus disease 2019 (COVID-19) was reported, and vaccines have been approved and rolled out in some countries to protect high-risk populations for severe outcomes, for the time being non-pharmaceutical interventions (NPIs) remain primary public health measures to slow the transmission and reduce the healthcare burden before vaccines are widely available and herd immunity can be achieved. Since the early stage of the pandemic, there have been a series of studies to understand the impacts of response policymaking and NPIs implementation on COVID-19, using mathematical modeling and simulations,¹⁻³ and different NPIs such as case isolation, close contact tracing and quarantine, social distancing, mask wearing, and travel restrictions have a varying effectiveness in controlling the transmission of COVID-19 across regions and time.⁴⁻¹⁰ However, few studies investigated the resurgences of COVID-19 transmission and it is little known about the effectiveness of NPIs for controlling secondary waves in regions where the COVID-19 epidemic has been contained.¹¹

The COVID-19 outbreak in Beijing's Xinfadi (XFD) Agricultural Wholesale market occurred on June 11, 2020 after 56 consecutive days without a newly confirmed case in Beijing, and the resurgence was brought under control in just 34 days from the onset of illness of the first known

case (June 5, 2020) to zero new infections detected (July 9, 2020).¹² It offers an exceptional opportunity to develop a model using real-world data and to quantitatively evaluate the timing and impact of integrated NPIs for containing COVID-19 resurgences. The source of this outbreak was identified as the virus spread from XFD market,¹³ the largest wholesale food market in Asia that has about 3,000 workers and 50,000 visitors each day and provides about 80% of Beijing's food supply. Within the massive XFD market complex, there are a total of 14 trading halls. One of which, the Beef and Mutton Trading Hall (BMTH), has been identified as the major infection source in this outbreak.¹³ After the new outbreak was discovered on June 11, the municipal government have adopted a two-pronged approach — re-instating NPIs used during the initial wave in January-March 2020 and introducing new NPIs including: i) large-scale tracing and management of high-risk populations identified by exposure risk levels, and ii) expanded SARS-CoV-2 nucleic acid screening in the general population in Beijing.

The implementation of combining interventions have rapidly and successfully contained the resurgence and interrupted the transmission, and only 362 confirmed cases, 40 asymptomatic infections in Beijing and 34 linked infections outside Beijing were found, with zero deaths and less disruption to routine socioeconomic activities.¹² To assess the effectiveness of the response to the COVID-19 outbreak in Beijing's XFD market and inform future response efforts of resurgence across regions, we developed a COVID-19 outbreak modeling framework to examine impacts of various identification timings and NPIs for this outbreak under hypothetical response scenarios.

Methods

Design. We constructed a modified susceptible–exposed–infectious–recovered (SEIR) model to evaluate the effectiveness of NPIs in containing COVID-19 after the outbreak in Beijing’s XFD market. We specifically modeled four scenarios for two key outcomes: the magnitude of transmission (ie, the number of cases) and the endpoint of transmission (ie, the date of containment). Our methods and results were reported according to guidelines developed by Bennett and Manuel.¹⁴ Ethical approval and informed consent requirements were waived by the Institutional Review Board and Human Research Ethics Committee of the Beijing Center for Disease Prevention and Control (Beijing CDC) because this study was considered a continuation of the public health investigation associated with an emerging infectious disease.

Data. The details of the Beijing XFD market outbreak, its investigation, and NPIs implemented in response, have been previously described.^{12,15} The data used in this study were extracted from the Notifiable Infectious Disease Reporting System (basic individual case-level demographics, location, and diagnostic data), the Epidemiological Investigation Information System (detailed individual case-level exposure, symptom, and clinical data), and the Close Contacts Tracing and Management System (individual contact-level demographics, exposures, location, and quarantine data).¹⁵ A summary of Beijing XFD outbreak data used in developing the model is presented in

Table 2.

The epidemiologic parameters such as incubation periods and contagious periods were calculated based on data from the epidemiological investigation and close contacts tracing and management. For incubation periods, we extracted 41 individual laboratory-confirmed records that have known dates of exposure. By combining the date of exposure with the date of symptom onset, we inferred the incubation periods for the 41 individual cases. We fitted the distribution of incubation periods to a Weibull distribution using a Maximum Likelihood Estimation method with R package

fitdistrplus.¹⁶ The contagious periods were calculated as the average duration from symptom onset to laboratory confirmation, since once the infections were confirmed, they would be quarantined and not cause a secondary infection. These and other parameters and coefficients used for model simulation are presented in **Table S1**.

Key populations in this model were defined as the assessed stratified risk groups by exposure level in the XFD outbreak. XFD workers who were in the XFD market were assessed to be at the highest risk. They were traced through traditional epidemiologic investigation methods (ie, face-to-face interviews or home visits) and quarantined in centralized facilities. Attack rate was calculated as the number of cases (numerator) divided by the number of total persons (denominator) presented as a percent. Attack rate among this high-risk group was 5.1%. The attack rate among workers in the BMTH was highest at 14.2%. Visitors to XFD market on June 12 were designated medium risk and quarantined in centralized facilities. They were found to have an attack rate of 0.1%. By contrast, visitors to XFD market before the outbreak (May 30–June 11) were assessed as low risk, traced by big data (ie, cell phone IP addresses), and were asked to quarantine at home. The attack rate among this low-risk group was 0.02% (**Table 1**). Close contacts were defined as persons who had direct contact within one meter with a confirmed case 4 days before or any time after their symptom onset without personal protective equipment. The close contact population was exclusive of, meaning did not meet the criteria for, the key population.

Model

Our model categorized the whole population into five subpopulations, susceptible (*S*), exposed and infected (but not yet infectious; *E*), infectious (*I*), infectious and isolated (*IS*) and removed (*R*). Based on the actual situation in the XFD outbreak, we further divided *E* population into a

subpopulation K indicating key population to the XFD market, a subpopulation C indicating close contacts of confirmed cases, a subpopulation N representing the infections detected by nucleic acid testing and a subpopulation O representing the infections identified by other methods. During this outbreak, a portion of the infected population was isolated early through close contact or key population tracing, or nucleic acid screening, such that it could not result in large scale secondary infection. Such infected population at the time of illness onset were classified as IS in our model. The IS group was further divided into two sub-populations according to the type of quarantine: those who were quarantined in centralized facilities (ISC) and those who were in home-based quarantine (ISH). The population R included the recovered (no deaths were reported during this outbreak) as well as individuals who were isolated or quarantined in designated facilities as they have no risk for secondary infections. Under those assumptions, we developed a modified SEIR model illustrated in **Figure 1**.

The equations are as follows:

$$\frac{dS_t}{dt} = -\beta S_t(I_t + IN_t) - q\beta S_t ISH_t, \quad (a)$$

$$\frac{dE_t}{dt} = \beta S_t(I_t + IN_t) + q\beta S_t ISH_t - \gamma E_t \quad (b)$$

$$\frac{dI_t}{dt} = (\gamma E_t - K_t - C_t - N_t) + (1 - \alpha_t^{1a} - \alpha_t^{1b} - \alpha_t^{1c} - \alpha_t^{1d} - \alpha_t^{1e} - \alpha_t^{1f})K_t - \sigma_{2t}I_t \quad (c)$$

$$\frac{dIN_t}{dt} = N_t - \sigma_{3t}IN_t \quad (d)$$

$$\frac{dISC_t}{dt} = (1 - h_t^2)C_t + (\alpha_t^{1a} + \alpha_t^{1b} + \alpha_t^{1c} + \alpha_t^{1d})K_t - \sigma_{1t}ISC_t \quad (e)$$

$$\frac{dISH_t}{dt} = h_t^2 C_t + (\alpha_t^{1e} + \alpha_t^{1f})K_t - \sigma_{1t}ISH_t \quad (f)$$

$$\frac{dR_t}{dt} = \sigma_{2t}I_t + \sigma_{3t}IN_t + \sigma_{1t}(ISC_t + ISH_t) \quad (g)$$

In the equation, K_t indicates the daily infections from the key population in XFD market, and $\alpha_t^{1a}, \alpha_t^{1b}, \alpha_t^{1c}, \alpha_t^{1d}, \alpha_t^{1e}$ and α_t^{1f} are the proportions of the daily infections in group 1.1 to group

3.2 (**Table 2**) of the total daily number in key population, respectively. $C_t = \alpha_2^t(\gamma E_t - K_t)$ indicates the daily infections discovered by close contact tracing, while parameter α_2^t models the effect of the close contact tracing and management. It is the proportion of the daily infections of the daily total number from close contacts, excluding the cases from the key population of the XFD market. $N_t = \alpha_3^t(\gamma E_t - K_t)$ indicates the daily infections found by expanded nucleic acid screening, while parameter α_3^t models the effect of the expanded nucleic acid screening. It is the infection proportion of daily total number from expanded nucleic acid screening, excluding the cases from the key population of the XFD market. Parameter σ_t represents the speed of transition from the infectious to the removed. Since all the infectious cases/infections were isolated in designated hospitals or home once confirmed, σ_t represents the speed of transition from the infectious to the confirmed. It equals the reciprocal of the time difference between the disease onset and confirmation. Parameter h_t^2 is the percentage of cases who were home-isolated among all the cases identified by closed contact tracing. These parameter values were abstracted from epidemiologic data (**Table 1, Table 2, Table S1**).

Parameter β is the transmission rate. Parameter q denotes the reduction of infectiousness for home-quarantined patients compared to non-quarantined patients. We estimated β and q by Markov Chain Monte Carlo (MCMC) with the Adaptive Metropolis algorithm implemented in the Python package *PyMC* (version 2.3.8).^{17,18} We used a non-informative flat prior of Uniform (9e-9, 5e-4) for β , and Uniform (0.01, 1.0) for q . We fitted the model with data of cumulative cases from June 6 to July 10, 2020, by the date of symptom onset. After a burn-in of 1,000 iterations, we ran the MCMC simulation for 30,000 iterations with a sampling size of each 10 iterations. The RMSE of cases from June 6 to July 10 between model prediction and observation is 16.61. All of these analyses were performed in Python (version 3.6.0).

Simulated Scenarios

Timing of initiation comprehensive measures – In the case of the XFD outbreak, the time between the first case being confirmed (June 11) and the XFD being closed (June 12) to contain the source of infection (i.e., infection source intervention) was 1 day.¹⁵ We hypothesized that delays in the implementation of the NPI, infection source intervention, would result in higher magnitude of transmission and later endpoint of transmission. Therefore, we modeled a 3-day, 7-day, and 14-day delay or ahead of the timing of infection source identification. We compared these predicted results with the actual outcomes from the XFD outbreak response: 368-case transmission magnitude and July 10 transmission endpoint. Assumptions made for this simulation included: (1) unlimited health resources, (2) all other NPIs implemented precisely as in the real outbreak response, and (3) the proportion of each group of key population in **Table 2** remains unchanged. We also changed the timing of other NPIs accordingly in model simulation (timing of the closure of XFD market, quarantining of key population, close contact tracing, expanded nucleic acid screening).

Management of key population – There were 224 confirmed infections identified in all managed key population. We simulated the epidemic development under quarantine protocols to the market workers or visitors to the market, respectively. In our simulation, for simplicity, we assumed that other NPI measures were independently unchanged when evaluating the effectiveness of each NPI measure of interest.

Quarantine of traced close contacts – There were 42 confirmed infections identified among close contacts group. For close contacts tracing, we compared the effect of centralized quarantine versus home quarantine.

Expanded screening for the general population – There were 28 confirmed infections identified from expanded screening in the general population. Expanded nucleic acid screening was expanded from the center of the XFD outbreak to surroundings among the general population. We simulated the effect of timing of nucleic acid test with 3 or 7 days delayed or in advance.

For all scenarios, we repeated the simulations based on parameter values estimated by 30,000 MCMC iterations with sampling at each 10 steps (i.e., 3,000 samples totally) to construct the 95% credible intervals of the epidemic curve by the 2.5 and 97.5 percentiles at each time point. The simulation results were presented as mean values and 95% credible intervals calculated from the 3,000 MCMC samples in this study. All the analyses were conducted using Python software, version 3.6.0.

Sharing The data used in this study are from public accessible and internal databases as well as news briefing indicated in the text; the mathematic model and code used for the analysis are available by addressing to the corresponding authors.

Results

Timing of combination NPI implementation

According to our model, a 3-day delay from the actual timing with which combination NPI response measures were initiated (i.e., identification of XFD as the infection source and initiation of NPIs) would lead to a 3-fold increase in the magnitude of transmission (95% CI: 2.6–3.4; ie, 1,104 versus 368 cases). A 7-day delay would lead to a 7.5-fold increase in the magnitude of transmission (CI: 6.4–8.6; ie, 2,768 versus 368 cases) and a 14-day delay a 28.2-fold increase (CI: 23.2–33.3; ie, 10,411 versus 368 cases). Finally, these delays would also lead to endpoints of transmission delayed by 31, 50, and 77 days, respectively (**Table 3, Figure 2 A and A'**).

Management of key populations

Similarly, if no quarantine was instituted, the magnitude of transmission would be 5.5-fold greater (CI: 4.5–6.2; i.e., 1,969 versus 368 cases) and the endpoint of transmission would be delayed 73 days compared to reality. If only the high- and medium-risk populations (ie, Group 1 and Group 2) were quarantined, the magnitude of transmission would be 1.7-fold greater (CI:1.5–2.0; 640 versus 368 cases) and the endpoint of transmission would be delayed 39 days (**Table 3, Figure 2 B and B'**). If just the targeted key population is modeled, no quarantine would lead to 8.2-fold greater (CI: 6.8–9.5; i.e., 1,825 versus 224 cases) magnitude of transmission and quarantine only for high- and medium-risk populations, a 2.2-fold greater (CI: 1.8–2.7; i.e., 496 versus 224 cases) magnitude of transmission (**Table 3**).

Quarantine of traced close contacts

According to our model, no quarantine for close contacts would result in 2-fold greater (CI: 1.7–2.3; i.e., 727 versus 368 cases) magnitude of transmission and a 44-day delay in the endpoint of transmission. A more rigorous all-centralized quarantine for close contacts would result in 1.0-fold (CI: 0.9–1.1; i.e., 361 versus 368 cases) the magnitude of transmission and a 5-day acceleration in the endpoint of transmission. By contrast, a less rigorous all-home quarantine for close contacts would result in 1.0-fold greater (CI: 0.9–1.2; i.e., 382 versus 368 cases) magnitude of transmission and a 4-day delay in the endpoint of transmission (**Table 3, and Figure 2 C and C'**). If just the targeted key population is modeled, no quarantine for close contacts results in 9.5-fold greater (CI: 6.7–12.3; 401 versus 42 cases) magnitude of transmission. More rigorous all-centralized quarantine for close contacts would result in 0.8-fold (CI: 0.1–1.6; 35 versus 42 cases) the magnitude of transmission whereas less rigorous all-home quarantine would result in 1.3-fold (CI: 0.3–2.3; 56 versus 42 cases) the magnitude of transmission (**Table 3**).

Expanded nucleic acid screening in the general population

According to our model, implementation of expanded nucleic acid screening 7 days earlier would result in 0.9-fold (CI: 0.8–1.0; 332 versus 368 cases) the magnitude of transmission and endpoint of transmission accelerated by 3 days. Three days earlier would result in 1.0-fold (CI: 0.9–1.0; 352 versus 368 cases) the magnitude of transmission and endpoint of transmission accelerated by 2 days. By contrast, a 3-day delay would result in 1.2-fold greater (CI: 1.0–1.3; 429 versus 368 cases) magnitude of transmission and an 8-day delay to the endpoint of transmission. A 7-day delay would result in 1.3-fold greater (CI: 1.1–1.5; 487 versus 368 cases) magnitude of transmission and a 17-day delay to the endpoint of transmission. Finally, no implementation of expanded nucleic acid screening at all would result in 1.6-fold greater (CI: 1.4–1.8; 603 versus 368 cases) magnitude of transmission and a 32-day delay to the endpoint of transmission (**Table 3, and Figure 2 D and**

D'). If just the targeted key population is modeled, a 3-day acceleration yields 0.4-fold (CI: 0.0–1.6; 12 versus 28 cases) the magnitude of transmission whereas a 3-day delay yields 3.2-fold greater (CI: 1.3–5; 89 versus 28 cases) magnitude of transmission, a 7-day delay results in a 5.3-fold greater (CI: 2.3–8.1; 147 versus 28 cases) magnitude of transmission, and no implementation at all yields a 9.4-fold greater (CI: 6.3–12.5; 263 versus 28 cases) magnitude of transmission (**Table 3**).

Discussion

Our results revealed that even just a 3 to 14-day delay in the identification of XFD as the infection source, subsequently leading a delay initiation of NPIs would have caused a 3 to 28-fold increase in transmission and a 31 to 77-day delay in the resurgence containment. A failure to implement the quarantine scheme employed in the XFD outbreak for at-risk groups would have caused 5-fold greater transmission and 73-day delay to containment. Similarly, failure to implement the quarantine plan executed in the XFD outbreak for close contacts would have caused 2-fold greater transmission and a 44-day delay to containment. Finally, failure to implement expanded nucleic acid screening would have yielded 1.6-fold greater transmission and a 32-day delay to containment.

Among the 272 infections found among the ~546 thousand individuals in defined key populations, 224 (82.4%) were identified through key population management while just 48 (17.6%) were identified through care seeking, contact tracing, expanded nucleic acid screening, or other methods. Our results show that failure to manage these key populations would have caused an 8-fold increase in the magnitude of transmission (ie, number of cases) and a 2-month delay in the endpoint of transmission (ie, containment) , indicating that this NPI is crucial. It is furthermore important

to note that this kind of intervention requires no sample collection, no laboratory testing, and no detailed epidemiological investigation. Rather, it is practical, straightforward, and highly targeted.

More recently, the cities of Wuhan and Qingdao in China have undertaken population-wide nucleic acid screening for SARS-CoV-2 among all residents in the cities, but this kind of action is costly and thus controversial.^{19,20} To be practical, real-time risk assessments must be paired with expanded nucleic acid screening so that efficiency is optimized. To control costs yet identify infections with maximum efficiency, the authorities in Beijing applied nucleic acid screening with real-time adjustment based on risk levels, expanding screening from the epicenter, XFD market, to surroundings gradually and ending it when the daily positivity rate fell to zero. Although our results seemed to indicate only a small 1.6-fold increase in transmission with failure to implement expanded nucleic acid screening, a 9.4-fold increase was predicted by our model for the key populations if no nucleic acid screening was conducted in the general population. Considering improved efficiency of expanded screening when coupled with ongoing risk assessment, this NPI should also be considered important for achieving containment of a COVID-19 outbreak.

Management of close contacts, whether in home or centralized quarantine, is also a controversial topic. Yet, close contact management has been proven to be effective.¹⁵ While home quarantine is more likely to cause secondary cases than centralized quarantine,¹⁵ centralized quarantine is much costlier. Moreover, centralized quarantine might trigger infection cluster, or even outbreak if the prevention and control measures are not effectively performed. Considering the cost of quarantine and the quality of life for quarantined persons, home-quarantine could be recommended if it could be conducted strictly. In fact, China's authority had modified the 14-day centralized quarantine

regulation to a 7-day centralized and 7-day home quarantine in the newly published protocol for COVID-19 prevention and control.¹⁶

This study has some limitations. In our simulations, we assumed that resources were adequate and all the NPIs were effectively implemented even as the numbers of cases surged. In reality, as disease transmission increased, some resources (e. g., centralized quarantine facilities) would be depleted and eventually reach a shortage. In addition, for simplicity, we assessed one NPI at a time, assuming the others would not be impacted. However, in reality, there are interdependencies between NPIs and the ability for officials to manipulate just one NPI alone at a time is limited. We also only evaluated two outcomes (i.e., magnitude of transmission and endpoint of transmission). Although these two are important with respect to containment, they fail to take other factors into account and ignore other outcomes of interest. Further study should be conducted using this model to evaluate the NPIs in different ways. For example, expanded nucleic acid screening in the general population appeared in our results to have only a small impact on the magnitude and endpoint of transmission. However, the extremely low positivity rate among the general population helped accelerate re-normalization and return to routine economic activity and social life in Beijing, limiting the negative consequences of anti-COVID-19 response activities. This socioeconomic benefit should be considered when evaluating NPIs in the future. Finally, social and culture factors also contributed to the success of containment efforts in the XFD market outbreak. Even though COVID-19 in China was at low level of transmission, and no new infections had been found in Beijing for nearly two months, residents were still wearing masks, obeying social distancing recommendations, and taking other personal prevention measures. They were also keenly aware of how dangerous SARS-CoV-2 infection was and how gravely other nations were being affected.

Thus, implementation of NPIs in the XFD outbreak was relatively well accepted, which may differ from circumstances in other settings in which this model may be applied.

Conclusions

In conclusion, this modeling study provides important new evidence that can not only inform the selection of NPI to use as countermeasures in response to a COVID-19 outbreak but also can inform the optimal timing of their implementation. Moreover, this modeling study clearly calculates the consequences of inaction and hesitation on the part of outbreak response teams and decision-making officials. This evidence should inform responses to future outbreaks of COVID-19 and future infectious disease outbreak preparedness efforts in China. Additionally, these lessons and this methodology can be used by other nations as they work to improve their anti-COVID-19 strategies and tactics.

List of abbreviations

Beijing CDC – Beijing Center for Disease Prevention and Control

BMTH – Beef and Mutton Trading Hall

CI – 95% credible interval

COVID-19 – coronavirus disease 2019

MCMC – Markov Chain Monte Carlo

NPI – non-pharmaceutical intervention

RMSE – root-mean-square error

SARS-CoV-2 – severe acute respiratory syndrome coronavirus 2

SEIR – susceptible–exposed–infectious–recovered

XFD – Xinfadi

Declarations

Ethics approval and consent to participate

Ethical approval and informed consent requirements were waived by the Institutional Review Board and Human Research Ethics Committee of the Beijing Center for Disease Prevention and Control (Beijing CDC) because this study was considered a continuation of the public health investigation associated with an emerging infectious disease.

Consent for publication

Not applicable.

Availability of data and materials

The data used in this study were extracted from the Notifiable Infectious Disease Reporting System (basic individual case-level demographics, location, and diagnostic data), the Epidemiological Investigation Information System (detailed individual case-level exposure,

symptom, and clinical data), and the Close Contacts Tracing and Management System (individual contact-level demographics, exposures, location, and quarantine data).¹⁵

Competing interests

The authors declare that they have no competing interests.

Funding

This study was funded by grants from Beijing Science and Technology Planning Project (Z201100005420010).

Authors' contributions

XW, ZW, QW, PY, and CH conceived the study and were in charge of overall direction and planning. XL, JW, XW, SL, ZJ, GL, and HS designed the modified SEIR model and performed the computations. XW, XH, JMM, KC, and SL wrote the manuscript. All authors discussed the results and contributed to and approved the final manuscript.

Acknowledgments

The authors would like to thank Chenxu Yu, Li Guo, Kangning Cai and Zhida Cheng from Yidu Cloud (Beijing) Technology Co., Ltd, for helping data collection and cleaning, manuscript revision and figure illustrations.

References

1. Brauner JM, Mindermann S, Sharma M, et al. Inferring the effectiveness of government interventions against COVID-19. *Science* 2020.

2. Lai S, Ruktanonchai NW, Zhou L, et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature* 2020; **585**(7825): 410-3.
3. Ruktanonchai NW, Floyd JR, Lai S, et al. Assessing the impact of coordinated COVID-19 exit strategies across Europe. *Science* 2020; **369**(6510): 1465-70.
4. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *Lancet* 2020; **395**(10242): 1973-87.
5. Fraser C, Riley S, Anderson RM, Ferguson NM. Factors that make an infectious disease outbreak controllable. *Proc Natl Acad Sci U S A* 2004; **101**(16): 6146-51.
6. Koo JR, Cook AR, Park M, et al. Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. *Lancet Infect Dis* 2020; **20**(6): 678-88.
7. Kretzschmar ME, Rozhnova G, Bootsma MCJ, van Boven M, van de Wijkert J, Bonten MJM. Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. *Lancet Public Health* 2020; **5**(8): e452-e9.
8. Leung K, Wu JT, Liu D, Leung GM. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. *Lancet* 2020; **395**(10233): 1382-93.
9. Tian H, Liu Y, Li Y, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* 2020; **368**(6491): 638-42.
10. Flaxman S, Mishra S, Gandy A, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* 2020; **584**(7820): 257-61.
11. Hao X, Cheng S, Wu D, Wu T, Lin X, Wang C. Reconstruction of the full transmission dynamics of COVID-19 in Wuhan. *Nature* 2020; **584**(7821): 420-4.

12. Wu Z, Wang Q, Zhao J, et al. Time Course of a Second Outbreak of COVID-19 in Beijing, China, June-July 2020. *Jama* 2020.
13. Pang X, Ren L, Wu S, et al. Cold-chain food contamination as the possible origin of COVID-19 resurgence in Beijing. *National Science Review* 2020; **7**(12): 1861-4.
14. Bennett C, Manuel DG. Reporting guidelines for modelling studies. *BMC Med Res Methodol* 2012; **12**: 168.
15. Zhou L, Wu Z, Li Z, et al. 100 Days of COVID-19 Prevention and Control in China. *Clin Infect Dis* 2020.
16. Delignette-Muller M, Dutang C. Fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical Software* 2015; **64**: 1-34.
17. Haario H, Saksman E, Tamminen J. An Adaptive Metropolis Algorithm. *Bernoulli* 2001; **7**.
18. Patil A, Huard D, Fonnesbeck CJ. PyMC: Bayesian Stochastic Modelling in Python. *2010* 2010; **35**(4): 81.
19. Cao S, Gan Y, Wang C, et al. Post-lockdown SARS-CoV-2 nucleic acid screening in nearly ten million residents of Wuhan, China. *Nat Commun* 2020; **11**(1): 5917.
20. Inocencio R. Coronavirus back in China: Entire city of Qingdao being tested after just 12 COVID-19 cases found. 2020. p. <https://www.cbsnews.com/news/coronavirus-china-qingdao-being-tested-12-covid-9-cases-found/>.
21. Bi Q, Wu Y, Mei S, et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *Lancet Infect Dis* 2020; **20**(8): 911-9.

22. Liang X, Feng Z, Li L. Guidance for Corona Virus Disease 2019:Prevention, Control, Diagnosis and Management: People's Medical Publishing House; WHO Collaborating Centre for Health Information and Publishing; 2020.

Figure 1. The illustration of the modified SEIR model.

Note: The model categorized the whole population into five subpopulations, susceptible (S), exposed and infected (not yet infectious) (E), infectious (I), infectious and isolated (IS) and removed (R). The IS group was further divided into two sub-populations according to the type of quarantine: those who were centralized quarantined (ISC) or home-based quarantined (ISH). E was further designated into a subpopulation K indicating key population to the XFD market, a subpopulation C indicating close contacts of confirmed cases, a subpopulation N representing the infections discovered by nucleic acid testing and a subpopulation O representing the infections identified by other methods.

Figure 2. Scenario model simulation to assess the timing of intervention, management of key populations, contact tracing and expanded SARS CoV-2 Nucleic Acid screening.

Note: **(A)** and **(A')** daily and cumulated total case number with the 3-day, 7-day, and 14-day delayed identification of XFD market as the source of infection. **(B)** and **(B')** daily and cumulated total case number simulated for management on the key population to the XFD market. **(C)** and **(C')** daily and cumulated total case number simulated for the close contacts traced with different quarantine protocols. **(D)** and **(D')** daily and cumulated total case number simulated for delayed or earlier expanded Nucleic acid screening on general population. Shading areas indicate 95% confidence intervals.

Table 1. Epidemiologic data collected for the XFD outbreak

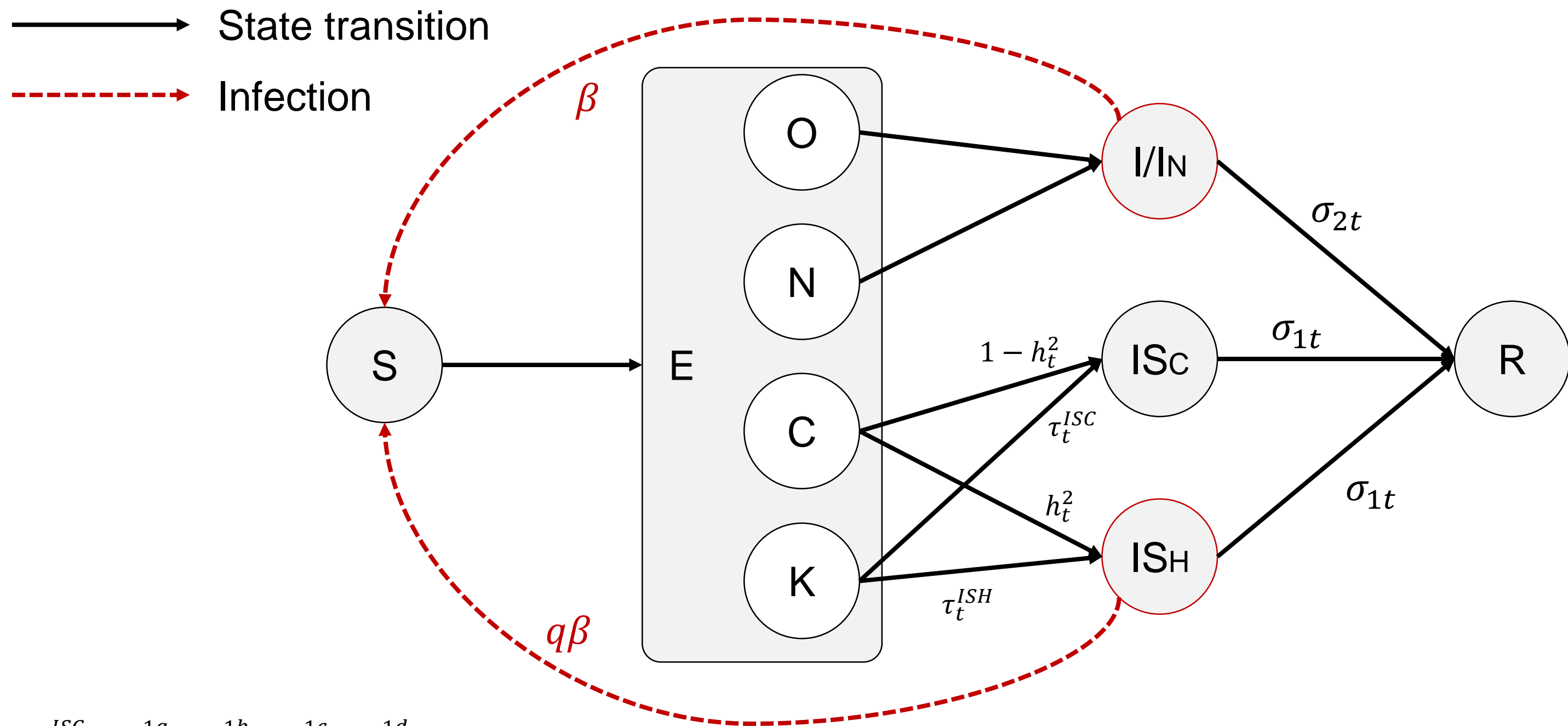
Categories	Case Numbers
Total infected cases during XFD outbreak	368
<i>Key Population</i>	
Number of individuals in key population	546,000
Number of Infections identified from key population	272
Number of Infections identified via key population management	224
<i>Close Contacts</i>	
Number of close contacts	6,607
Number of close contacts with home-based quarantine	1,055
Number of close contacts with centralized quarantine	5,552
Number of infected close contacts	84
Number of Infected cases identified via contact tracing	42
<i>Expanded Nucleic Acid Screening</i>	
Number of people conducted nucleic acid screening	10,878,289
Number of Infected cases identified via expanded nucleic acid screening	28
<i>Other methods</i>	
Number of target population	10,111,711
Infected cases identified	76

Table 2. Group of key population and attack rate of each group

Group	Exposed populations	Quarantine protocol	Tracing techniques	Risk level	Total No.	No. of cases	Attack rate (%)
1	Workers at the XFD market	Centralized quarantine	Onsite	High	3,311	169	5.10
1.1	Workers at the BMTH	Centralized quarantine	Onsite	High	838	119	14.20
1.2	Workers at the XFD market other than BMTH	Centralized quarantine	Onsite	High	2,473	50	2.02
2	Visitors to the XFD market on June 12 th	Centralized quarantine	Onsite	Medium	7,689	8	0.10
2.1	Visitors to the BMTH market	Centralized quarantine	Onsite	Medium	1,078	8	0.74
2.2	Visitors to the XFD market other than BMTH	Centralized quarantine	Onsite	Medium	6,611	0	0
3	Visitors to XFD between May 30 th and June 11 th	Home quarantine	Big data	Low	535,000	95	0.02
3.1	Visitors to the BMTH market	Home quarantine	Big data	Low	75,000	64	0.09
3.2	Visitors to the XFD market other than BMTH	Home quarantine	Big data	Low	460,000	31	0.007

Table 3. Results of scenario simulations on two outcomes: magnitude of transmission and endpoint of transmission.

Scenarios Simulated	Magnitude of Transmission			Endpoint of Transmission	
	Number of Cases (95% CI)	Relative Increase Fold (95% CI)	Number of Cases in target-population (95% CI)	Relative Increase Fold (95% CI)	Date of Containment (Relative Delay)
<i>Timing of initiation comprehensive measures</i>					
Reality: Delayed 0 days	368	–	368		July 10
Scenario 1: Delayed 3 days	1,104 (947-1,261)	3.0 (2.6–3.4)	1,104 (947-1,261)	3.0 (2.6–3.4)	August 10 (31 days)
Scenario 2: Delayed 7 days	2,768 (2,360-3,176)	7.5 (6.4–8.6)	2,768 (2,360-3,176)	7.5 (6.4–8.6)	August 29 (50 days)
Scenario 3: Delayed 14 days	10,411 (8,549-12,272)	28.2 (23.2–33.3)	10,411 (8,549-12,272)	28.2 (23.2–33.3)	September 25 (77 days)
<i>Management of Key Population</i>					
Reality:	368	–	224		July 10
Scenario 1: No quarantine	1,969 (1,658-2,280)	5.5 (4.5-6.2)	1,825 (1,514-2,136)	8.2 (6.8-9.5)	September 21 (73 days)
Scenario 2: Quarantine high-medium risk	640 (536-744)	1.7 (1.5-2.0)	496 (392-600)	2.2 (1.8-2.7)	August 18 (39 days)
<i>Quarantine of traced close contacts</i>					
Reality: Centralized and home quarantine	368		42		July 10
Scenario 1: No quarantine	727 (609-844)	2.0 (1.7-2.3)	401 (283-518)	9.5 (6.7-12.3)	August 23 (44 days)
Scenario 2: Centralized quarantine for all	361 (330-392)	1.0 (0.9-1.1)	35 (4-66)	0.8 (0.1-1.6)	July 5 (-5 days)
Scenario 3: Home quarantine for all	382 (341-423)	1.0 (0.9-1.2)	56 (15-97)	1.3 (0.3-2.3)	July 14 (4 days)
<i>Expanded nucleic acid screening in the general population</i>					
Reality:	368		28		
Scenario 1: Accelerated 7 days	332 (305-358)	0.9 (0.8-1.0)	-8 (-35-18)	0	July 7 (-3 days)
Scenario 2: Accelerated 3 days	352 (319-384)	1.0 (0.9-1.0)	12 (-21-44)	0.4 (0-1.6)	July 8 (-2 days)
Scenario 3: Delayed 3 days	429 (377 to 481)	1.2 (1.0-1.3)	89 (37-141)	3.2 (1.3-5)	July 18 (8 days)
Scenario 4: Delayed 7 days	487 (405 to 568)	1.3 (1.1-1.5)	147 (65-228)	5.3 (2.3-8.1)	July 27 (17 days)
Scenario 5: No Nucleic Acid Screening	603 (516 to 690)	1.6 (1.4-1.8)	263 (176-350)	9.4 (6.3-12.5)	August 11 (32 days)



$$\tau_t^{ISC} = \alpha_t^{1a} + \alpha_t^{1b} + \alpha_t^{1c} + \alpha_t^{1d}$$

$$\tau_t^{ISH} = \alpha_t^{1e} + \alpha_t^{1f}$$

