Transmission dynamics, heterogeneity, and controllability of SARS-CoV-2 in the rural area

Yuying Li  
Peking University  https://orcid.org/0000-0002-3231-7542

Taojun Hu  
Peking University

Xin Gai  
Peking University

Yunjun Zhang  
Peking University

Xiao-Hua Zhou  (azhou@math.pku.edu.cn)  
Peking University  https://orcid.org/0000-0001-7935-1222

Research Article

Keywords: SARS-CoV-2, urban–rural, heterogeneity, vaccination, non-pharmaceutical interventions

Posted Date: April 6th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-360060/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Transmission dynamics, heterogeneity, and controllability of SARS-CoV-2 in the rural area

Yuying Li¹, Taojun Hu¹, Xin Gai¹, Yunjun Zhang¹, Xiaohua Zhou¹,²,³*

¹ Department of Biostatistics, School of Public Health, Peking University, Beijing 100191, China

² Beijing International Center for Mathematical Research, Peking University, Beijing 100871, China

³ Center for Statistical Sciences, Peking University, Beijing 100871, China

* Correspondence: azhou@math.pku.edu.cn
ABSTRACT

Backgrounds: Few studies examine the transmission dynamics and heterogeneity of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in rural areas and clarify rural–urban differences. Moreover, the effectiveness of non-pharmaceutical interventions (NPIs) relative to that of vaccination in rural areas is uncertain.

Methods: We addressed this knowledge gap using an improved statistical stochastic method based on the Galton–Watson branching process considering both symptomatic and asymptomatic cases. Data were collected from the epidemiological records of 1136 SARS-2-CoV infections after the rural outbreak in Hebei, China, between 2 January and 20 February 2021.

Results: The estimated average reproductive number $R$ and dispersion parameter $k$ ($k<1$ indicating strong heterogeneity) in the rural area were 0.55 (95% confidence interval [CI]: 0.45–0.68) and 0.14 (95% CI: 0.10–0.20), respectively. Although age group and contact-type distributions significantly differed between urban and rural areas, the $R$ and $k$ did not. Further, simulation results based on pre-control parameters ($R = 0.81, k = 0.27$) showed that in the vaccination scenario (80% efficacy and 55% coverage), the cumulative secondary infections will be reduced by more than half; however, NPIs are more effective than vaccinating 65% of the population. The presence of asymptomatic infections might affect the estimation of $R$ but showed no significant effect on estimating transmission heterogeneity.

Conclusion: The government should pay equal value to SARS-CoV-2 transmission in rural and urban areas and conduct specific prevention and control measures in rural areas. Older adults and children should receive particular attention in such policies, and
community contact should be minimized. Moreover, the government must consider a country’s economy and governance when conducting vaccination and NPIs.

**Keywords:** SARS-CoV-2, urban–rural, heterogeneity, vaccination, non-pharmaceutical interventions

**INTRODUCTION**

The dynamics of an outbreak depend on the average reproductive number \( R \) and individual heterogeneity in transmission. Although there is ample research on \( R \), studies regarding heterogeneity are limited. Heterogeneity reflects the divergence of secondary infections in each case and can be estimated by describing the distribution of secondary cases as a negative binomial distribution with dispersion parameter \( k \), where \( k < 1 \) suggests that the transmission is over-dispersed[1]. Diseases with high heterogeneity show infrequent but explosive epidemics; for example, in 2003, many settings experienced no epidemic despite unprotected exposure to severe acute respiratory syndrome (SARS) cases[2,3], whereas a few cities suffered explosive outbreaks of SARS [4,5]. Thus, understanding the role of transmission heterogeneity in severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) dynamics is important for outbreak control.

However, few studies explore the impact of asymptomatic infection on disease dynamics, especially individual heterogeneity in transmission. This is likely due to the lack of a valid statistical model and ambiguity of fundamental epidemiological questions that remain poorly understood, such as the proportion of asymptomatic cases[6]. Nishiura et al estimated that the asymptomatic ratio of coronavirus disease (COVID-19) was
41.6% (5 out of 12 confirmed cases) among 565 Japanese individuals evacuated from Wuhan, China [7]. A recent review [8] of 41 studies showed that the pooled percentage of asymptomatic infection was 15.6% (95% confidence interval [CI]: 10.1–23.0%).

Moreover, most studies focus solely on urban areas and ignore rural regions. Rural regions tend to have higher levels of poverty [9] and fewer job opportunities [10] relative to urban areas. Furthermore, rural areas broadly lack access to healthcare [11], tend to have older and healthier populations [12,13], and lack awareness of timely medical treatment [14]. The gaps lead to discrepancy in transmission dynamics between urban–rural areas, warranting improved corresponding control policies.

The second-wave outbreak in Hebei, China, mainly occurred in rural regions and swiftly subdued. After the first confirmation on 2 January 2021, the government implemented city-wide nucleic acid tests (NAT) in the two most severely affected cities to detect symptomatic and asymptomatic infections. Concurrently, highly detailed epidemiological information on individuals and their close contacts was collected by the Health Commission of Hebei province.

In this study, we provide an evidence-based picture of the SARS-CoV-2 outbreak in rural areas, where the government should pay more attention to older adults, children, and community contacts when conducting prevention and control measures. Additionally, we extended a statistical model that can be applied to other regions and provide more comprehensive results considering symptomatic and asymptomatic cases. To the best of our knowledge, this is the first direct comparison of the effects of non-pharmaceutical interventions (NPIs) under actual conditions and vaccination.
METHODS

Data collection

We collected detailed data on 942 confirmed SARS-CoV-2 infections in Hebei province, China, from 2 January to 20 February 2021, which were available in the website of the Health Commission of Hebei province [15]. In addition, asymptomatic infections accounted for 17% (194/1136) of the total infections. No new infections have been confirmed in Hebei province since February 14, indicating that the outbreak has been controlled. Primary and secondary SARS-CoV-2 infections were identified through: (i) active screening of incoming passengers in Hebei province, especially those who had travelled to areas defined by the Chinese government [16] as medium- or high-risk to capture travel-associated symptomatic and asymptomatic infections; (ii) passive surveillance in hospitals and outpatient clinics, involving testing of individuals suspected with COVID-19 to capture symptomatic cases; (iii) contact tracing of all confirmed infections identified by the above screening, followed by systematic monitoring of their close contacts, to capture symptomatic and asymptomatic infections; and iv) city-wide NAT to capture symptomatic and asymptomatic infections. From 6 January to 22 January 2021, Shijiazhuang City, the most severely affected area in Hebei province, has carried out three rounds of full-staff NAT, with the total exceeding 30 million person-times.

The collected data for each confirmed case included age, sex, prefecture, date of symptom onset, date of diagnosis of asymptomatic infection, date of confirmation, potential exposures, and contact history. Based on contact tracing, a SARS-CoV-2 cluster was defined as a group of $\geq 2$ confirmed SARS-CoV-2 cases or asymptomatic infections with an epidemiological link, i.e. occurring through the same contact-type (e.g., home,
social, community, or other). According to the extent of resolution of the reconstructed infection cluster, i.e. the number of primary cases and chain size (total number of cases in the transmission chain), three chain types were further identified (Simple/Ordinary/Complex). We considered the sporadic cases as isolated simple transmission chains (detailed definitions are provided Supplement, Section 1). Because contact tracing of asymptomatic infections was unavailable, we imputed them into the whole transmission chain according to the rates of secondary cases (imputation mechanisms are provided in Supplement, Section 2).

We additionally collected detailed epidemiological records for the outbreak of 135 cases confirmed in Tianjin [17] from 21 January 2020 to 26 February 2020. We selected Tianjin, a municipal city in China, as a representative city of the urban outbreak and compared the differences in transmission dynamics, demographics, and serial intervals between urban and rural areas. The definition of rural and urban areas in China according to relative regulations[18].

**Statistical analysis**

**Inference about Transmission Characteristics**

To estimate the $R$ and $k$ simultaneously, we deployed a statistical stochastic method based on the Galton–Watson branching process to simulate the entire transmission. Assuming that the offspring distribution follows a negative binomial distribution with mean $R$ and dispersion parameter $k$ (lower $k$ indicates higher heterogeneity), we estimated the two parameters using maximum likelihood estimation (MLE) [19,20]. Additionally, by considering asymptomatic cases, we improved their method to fit into a
wider population including symptomatic and asymptomatic cases. To guarantee the robustness of the estimation results, we adopted two approaches to infer the corresponding CIs of \( R \) and \( k \), namely, the LRT and the biased-corrected and accelerated bootstrap methods. The model details and parameter settings are provided in Sections 3 and 4 of the Supplement.

**Estimation of serial interval**

To compare the difference between the serial intervals of SARS-CoV-2 in rural and urban areas, we fitted serial intervals retrieved from a total of 8 and 12 infector–infectee pairs in Hebei and Tianjin into a parametric Weibull model. Statistically significant differences were determined by conducting the LRT.

**Assessment of different interventions**

We analyzed the effectiveness of two types of interventions on SARS-CoV-2 transmission: vaccination and NPIs, including city-wide NATs, isolation, and mask-wearing. For the NPIs, we divided the entire outbreak period into three segments according to the time of two rounds of city-wide NAT (9 January and 14 January 2021), and estimated \( R \) and \( k \) for each time period. We used the parameters before the first NAT intervention as the baseline and the parameters for the second period (during 9 January and 14 January 2020) as the effect of non-pharmaceutical intervention. To clarify the impact of the vaccine and its comparison with NPIs, we assumed that the population had been vaccinated before the outbreak (80\% efficacy) and simulated the cumulative secondary infections, considering a range of coverage rates (20\%, 55\%, 65\%, and 75\%). All interventions were compared with the baseline (without control measures).
RESULTS

Characterizing SARS-CoV-2 transmission chains and heterogeneity in rural areas

In total, 942 SARS-CoV-2 confirmed infections, except for two infections with missing information, occurred in 387 (41.1%) males and 553 (58.8%) females. We observed 0–4 generations of transmission, with the largest cluster involving 44 SARS-CoV-2–infected individuals. Exposures were grouped into four categories according to contact type: household, community, social, and primary case (Definitions in Section 5 of Supplement). Except for the primary case, household contacts accounted for the highest proportion of transmission, followed by community contacts and social contacts. Figure 1 shows the reconstructed transmission chain.
Figure 1. SARS-CoV-2 transmission chains. Reconstructed transmission chains of 942 SARS-CoV-2-infected individuals in the rural outbreak in Hebei province. Each node in the network represents a patient infected with SARS-CoV-2, and each link represents an infector–infectee relationship. The color of the node denotes the reporting contact type of the infected individuals. The size of the node corresponds to the number of secondary cases.

Table 1. Characteristics of the three types of transmission chains for SARS-CoV-2 outbreak in rural areas

| Chain type | Total number of chains (n=655) | Total number of cases (n=942) | Average chain size | Range of chain size |
|------------|-------------------------------|-------------------------------|--------------------|---------------------|
| Simple     | 639                           | 751                           | 1.2                | 1–5                 |
| Ordinary   | 15                            | 186                           | 12.4               | 3–44                |
| Complex    | 1                             | 4                             | 4                  | 4–4                 |
SARS-CoV-2, severe acute respiratory syndrome coronavirus 2

The 942 SARS-CoV-2-confirmed infections were grouped into 655 transmission chains, including 639 simple chains (752 infections, average size: 1.2), 15 ordinary chains (186 infections, average size: 12.4), and one complex chain (4 infections, average size: 4; Table 1). Considering asymptomatic infections, the average $R$ of outbreak in the rural area was 0.55 (Table 2); this represents a low transmission risk. The 95% CIs of $R$ estimated by the likelihood radio test (LRT) and biased-corrected and accelerated bootstrap (BCa bootstrap) methods were 0.45–0.68 and 0.44–0.69, respectively. There was no significant difference between the results obtained by the two methods, indicating the robustness of our results. The dispersion parameter $k$ was 0.14, and 95% CIs were 0.10–0.20 and 0.10–0.19 estimated by LRT and BCa bootstrap, indicating considerable heterogeneity of SARS-CoV-2 transmission in the rural area.

**Table 2.** Estimation of average reproductive number ($R$) and dispersion parameter ($k$).

| Average reproductive number ($R$) | Dispersion parameter ($k$) |
|-----------------------------------|----------------------------|
| Point estimation                  | 0.55                       | 0.14                       |
| 95% CI by LRT                     | (0.45, 0.68)               | (0.10, 0.20)               |
| 95% CI by BCa Bootstrap           | (0.44, 0.69)               | (0.10, 0.19)               |

CI, confidence interval; LRT, likelihood radio test; BCa Bootstrap, biased-corrected and accelerated bootstrap

**Comparison of SARS-CoV-2 transmission characteristics between the urban and rural areas**

There were no significant differences in $R$ and $k$ between the two areas ($p < 0.05$, Table 3). The $R$ values were 0.74 and 0.55 and $k$ values were 0.35 and 0.14 for urban and
rural areas, respectively. However, there were significant differences in age, sex, and contact-type distributions between these two areas. The proportions of older adults (≥ 65 years old) and children (≤ 20 years old) in the rural area were higher than those in the urban area (16.1% vs. 14.8% and 16.4% vs. 3.7%, respectively). Although more than 50% of transmissions in both urban and rural areas were caused by household contacts (61.1% and 51.4%), community contacts also accounted for a large proportion (46.5%) in rural areas. The median serial interval of the rural area was shorter than that of the urban area (5.5 days vs. 6.0 days), although without significant differences.

Table 3. Comparison of SARS-CoV-2 transmission between urban and rural areas

|                          | Urban area (Tianjin, n=135) | Rural area (Hebei, n=942) | P value |
|--------------------------|-----------------------------|----------------------------|---------|
| Age, years               |                             |                            | <0.001* |
| Median (IQR)             | 49 (36–62)                  | 46 (30–60)                 |         |
| <20                      | 5 (3.7%)                    | 155 (16.4%)                |         |
| 20–64                    | 110 (81.4%)                 | 633 (67.2%)                |         |
| ≥65                      | 20 (14.8%)                  | 152 (16.1%)                |         |
| Sex                      |                             |                            | 0.01*   |
| Female                   | 63 (46.7%)                  | 553 (58.7%)                |         |
| Male                     | 72 (53.3%)                  | 387 (41.1%)                |         |
| Contact type             |                             |                            | <0.001* |
| Household                | 55 (61.1%)                  | 147 (51.4%)                |         |
| Social                   | 15 (16.7%)                  | 6 (2.1%)                   |         |
| Community                | 20 (22.2%)                  | 133 (46.5%)                |         |
| Median of serial interval| 5.5 (IQR: 3.6–7.8)          | 6.0 (IQR: 3.6–9.0)         | 0.73    |
Transmission dynamics

|   | R (95% CI)                  | k (95% CI)                  |
|---|----------------------------|----------------------------|
|   | 0.74 (0.51–1.10)           | 0.16                       |
|   | 0.55 (0.45–0.68)           | 0.09                       |
|   | 0.35 (0.13–1.21)           | 0.14                       |
|   | 0.14 (0.10–0.20)           | 0.09                       |

Student’s t-test was used to compare the differences in age groups. The $\chi^2$ test was used to compare differences in sex and contact type. LRT was used to compare the difference in the serial interval and transmission dynamics. The estimation of $R$ and $k$ is based on the imputed dataset with a total of 1136 infections. IQR, interquartile range.

**Assessment of NPIs and vaccination in SARS-CoV-2 transmission**

For NPIs, until the first NAT, the average $R$ was 0.81 (95% CI: 0.65–1.02) and dispersion parameter $k$ was 0.27 (95% CI: 0.14–0.56; Table 4). After the first NAT and before the second NAT, $R$ decreased significantly to 0.33 (95% CI: 0.22–0.50) and remained around this level after the second NAT, while the $k$ value showed a downward trend at first but then increased slightly to 0.17 (95% CI: 0.10–0.31) without significant changes. We concluded that the first round of NAT played the most significant role in curbing the spread of SARS-CoV-2.

**Table 4. Impact of non-pharmaceutical interventions on SARS-CoV-2 transmission dynamics**

|                                    | $R$ (95% CI) | $k$ (95% CI) |
|------------------------------------|--------------|--------------|
| Before 1st round citywide-NAT (<1/09) | 0.81 (0.65,1.02) | 0.27 (0.14,0.56) |
| During 1st to 2nd round citywide-NAT (1/09-1/14) | 0.33 (0.22,0.50) | 0.13 (0.07,0.23) |
| After 2nd round citywide-NAT (>1/14) | 0.36 (0.25,0.55) | 0.17 (0.10,0.31) |
We conducted simulation studies with 656 primary cases and estimated the cumulative offspring infections from 1 January to 31 March. In Figure 2(a), the cumulative offspring infections in the rural area were projected to reach 1575 (95% CI: 1110–2094) under the baseline setting ($R = 0.81$, $k = 0.27$). In other words, the total number of infections of the outbreak in the rural outbreak will reach 2231, as the number of primary cases is 656. Comparing the final secondary size in vaccinating different proportions of the population, with 55% of the population vaccinated, the cumulative offspring infections reduced by 51% at 769 (95% CI: 468–1156). In 75% of the population vaccinated, the cumulative offspring infections reduced by 70% at 478 (95% CI: 272–789). As shown in Figure 2(b), the final secondary infections were 264 (95% CI:186–353), showing that the number of infected people reduced by approximately 83%. Considering the effectiveness of interventions, the effect of vaccinating all personnel was better than that of the current NPIs implemented in rural areas, while current control is superior to vaccinating 65% of the population.
(a) Impact of vaccination (under different levels of the vaccinated population: 20%, 55%, 65%, and 75%) interventions

(b) Comparison of vaccination and non-pharmaceutical interventions

Figure 2. Comparison of NPIs and vaccination in SARS-CoV-2 transmission.

Sensitivity analyses

In addition, the robustness of the model was verified. At first, we identified the role of asymptomatic infections in the epidemic dynamics. Given the higher proportion of asymptomatic infections, the estimates for $R$ showed a steadily increasing trend, from 0.51 to 1.95. The estimates of $k$ fluctuated around 0.14. This indicated that with a higher estimation of the proportion of asymptomatic cases, based on a parallel dataset, the estimated $R$ was larger, indicating a more severe potential spread of SARS-CoV-2. In contrast, there was no significant change in the heterogeneity of disease transmission due to the proportion of asymptomatic infections. Moreover, we evaluated the performance of our model based on the idea of 10-fold cross-validation[21]. The data were divided into 10 equal parts, and nine of them were randomly selected for parameter estimation. The mean values (standard deviations) of $R$ and $k$ were 0.55 (0.02) and 0.14 (0.01), respectively, verifying the good performance of our model when generalized to an independent dataset (Table 5).

Table 5. Sensitivity analyses. The impact of asymptomatic infections and cross-validation

| Infections            | $R$    | $k$    |
|----------------------|--------|--------|
| Null asymptomatic    | 0.51   | 0.13   |
| infections (p=0)     |        |        |
| 20% asymptomatic     | 0.56   | 0.14   |
| infections (p=0.2)   |        |        |
| 40% asymptomatic     | 0.63   | 0.17   |
| infections (p=0.4)   |        |        |


| Asymptomatic Infections | $R$ | SD |
|-------------------------|-----|----|
| 60% asymptomatic infections (p=0.6) | 1.62 | 0.09 |
| 80% asymptomatic infections (p=0.8) | 1.78 | 0.10 |
| All asymptomatic infections (p=1) | 1.95 | 0.13 |
| Mean (SD) for 10-fold cross-validation | 0.55 (0.02) | 0.14 (0.01) |

$k$, dispersion parameter; $R$, average reproductive number; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2

**DISCUSSION**

We characterized the transmissibility of SARS-CoV-2 in rural areas with the presence of asymptomatic infections based on the detailed epidemiology records in Hebei, China, and further compared the effectiveness of vaccination and that of NPIs. SARS-CoV-2 transmission in the urban and rural areas showed a strong heterogeneity. Moreover, household contact was the most important mode of transmission, whether in the city or the countryside, but community contact also played an important role in countryside transmission. We also found that in the vaccination scenario (80% efficacy and 55% coverage), the cumulative secondary infections will be reduced by more than half; however, NPIs are more effective than vaccinating 65% of the population. The presence of asymptomatic infections might affect the estimation of $R$ but showed no significant effect on estimating transmission heterogeneity.

The estimated dispersion parameter $k$ was 0.14 (95% CI: 0.10–0.20) in the rural area, indicating strong transmission heterogeneity. This result is consistent with that of another study. Lau et al [22] reported that a rural area (Dougherty) in Georgia, USA, had strong transmission heterogeneity ($k = 0.43; 95\% \text{ CI: } 0.39–0.47$). Although there was no significant urban–rural difference in $k$ under 5% type I error ($P=0.09$), $k$ in the rural area was lower than that in the urban area, in concordance with the results of Lau et al.
Transmission heterogeneity results from many factors including pathogen virulence, control measures, and activity density. The results of whole-genome sequencing and phylogenetic analysis revealed that the strains of Tianjin and Hebei both belonged to the European branch of the lineage (L-Lineage)\[23,24\]. However, large-scale NAT, which was not carried out in Tianjin, can quickly and comprehensively screen asymptomatic and mild symptomatic infections, thereby effectively shortening the infection time. Additionally, the rural outbreak in Hebei coincided with China’s spring festival and several weddings, thus, increasing the probability of large gatherings. Therefore, the government should pay sufficient attention to rural areas instead of focusing solely on urban areas.

SARS-CoV-2 transmission in rural areas had similar transmission dynamics as that in urban areas but differed in terms of age group and contact-type distributions. Our results indicated that rural areas had a larger proportion of older adults and children with COVID-19 than did the urban areas. This may be because rural areas had older populations, on average, and more people with underlying health conditions than suburban and urban communities \[12\]. Additionally, older adults are more likely to be hospitalized and have severe COVID-19, with higher mortality rates \[25\]. Household transmission played an important role in both outbreaks, which corroborates previous studies \[26,27\]. However, in rural areas, community contacts also lead to a large proportion of infections. Similar to that in other rural areas \[28\], most community contacts in Hebei were consequent to wedding receptions. Therefore, the government must develop prevention and control measures in rural areas, mainly focusing on older
adults and children and restricting large gatherings that pose a high risk for infectious disease transmission [25].

We found that NPIs lead to a larger reduction in infections than vaccination (80% efficacy and 65% coverage). This may explain why the outbreak was swiftly controlled in Hebei. A recent study also verified that NPIs are cost-effective approaches to curb the spread of SARS-CoV-2 [29]. However, the effect of NPIs is closely related to the timing and quality of implementation; hence, similar strategies might have different effects in different cities [30]. Therefore, countries with strong governance can prioritize NPIs until vaccines are widely available. Nonetheless, the durability of responses after vaccination is uncertain [31]; therefore, vaccination is more suitable for countries with economic strength but weak governance.

Our findings have several limitations: First, contact tracing of asymptomatic infections is not provided, which may give rise to biased reconstructions of transmission chains. However, detailed records of asymptomatic infections are difficult to collect because they require intensive prospective clinical sampling and screening, which hinders many studies. We imputed this information by assuming missing at random, and our sensitivity analyses can prove the accuracy of our model to a certain extent, but more detailed epidemic data would be helpful. Second, we explicitly set the fixed value of the proportion of asymptomatic infections when estimating the model parameters. Although we conducted sensitivity analyses, improved methodology and more accurate estimation of the asymptomatic proportion is a future research direction. Lastly, the regulation of asymptomatic infections on transmission dynamics needs to be further explored and evidenced.
CONCLUSION

The government should pay equal value to SARS-CoV-2 transmission in rural and urban areas and conduct specific prevention and control measures in rural areas. Older adults and children should receive particular attention in such policies, and community contact should be minimized. Moreover, the government must consider a country’s economy and governance when conducting vaccination and NPIs.

Declarations

Ethical Approval and Consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The code to perform statistical analysis is available from the authors upon request. The data of the outbreak in Hebei and Tian are available in this paper and from the authors upon request.

Competing interests

The authors declare that they have no competing interests

Funding

This work was supported by grant from the National Natural Science Foundation of China (Grant number: 82041023).

Author contributions
X.Z., Y.Z., and Y.L. conceived the study and led the over-all scientific questions. Y.L., T.H. and X.G. carried out data analysis and modelling studies. Y.L. collected and processed the data. Y.L. and T.H. wrote the paper. Y.L., T.H., and X.G. revised the paper.

**Acknowledgements**

We thank the health workers who contributed to the epidemiological survey. We thank researchers at Peking university for their contributions on data collection.

**References**

1. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. Nature. 2005;438:355–9.

2. Park BJ, Peck AJ, Kuehnert MJ, Newbern C, Smelser C, Comer JA, et al. Lack of SARS Transmission among Healthcare Workers, United States. Emerg Infect Dis. 2004;10:244–8.

3. Ha LD, Bloom SA, Hien NQ, Maloney SA, Mai LQ, Leitmeyer KC, et al. Lack of SARS Transmission among Public Hospital Workers, Vietnam. Emerg Infect Dis. 2004;10:265–8.

4. Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, et al. Transmission dynamics and control of severe acute respiratory syndrome. Science. 2003;300:1966–70.
5. Riley S, Fraser C, Donnelly CA, Ghani AC, Abu-Raddad LJ, Hedley AJ, et al. 
Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of 
public health interventions. Science (80- ). 2003;300:1961–6.

6. Subramanian R, He Q, Pascual M. Quantifying asymptomatic infection and 
transmission of COVID-19 in New York City using observed cases, serology and 
testing capacity. medRxiv. 2020;1–10.

7. Nishiura H, Kobayashi T, Miyama T, Suzuki A, Jung S-M, Hayashi K, et al. 
Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). 
Int. J. Infect. Dis. IJID Off. Publ. Int. Soc. Infect. Dis. 2020. p. 154–5.

8. He J, Guo Y, Mao R, Zhang J. Proportion of asymptomatic coronavirus disease 
2019: A systematic review and meta-analysis. J Med Virol. 2021;93:820–30.

9. Tickamyer AR, Sherman J, Warlick J, editors. Rural Poverty in the United States 
[Internet]. Columbia University Press; 2017. Available from: 
http://www.jstor.org/stable/10.7312/tick17222

10. Green GP. 15. The Opportunities and Limits of Economic Growth. Rural poverty United States. Columbia University Press; 2017. p. 416–38.

11. Bailey C, Jensen L, Ransom E. Rural America in a Globalizing World. West Virginia University Press; 2014.

12. Henning-Smith C. The Unique Impact of COVID-19 on Older Adults in Rural Areas. J Aging Soc Policy. England; 2020;32:396–402.
13. Peters DJ. Community Susceptibility and Resiliency to COVID-19 Across the Rural-Urban Continuum in the United States. J Rural Heal Off J Am Rural Heal Assoc Natl Rural Heal Care Assoc. 2020;36:446–56.

14. Lei L, Richards JS, Li Z-H, Gong Y-F, Zhang S-Z, Xiao N. A framework for assessing local transmission risk of imported malaria cases. Infect Dis poverty. 2019;8:43.

15. Health Commision of Hebei province [Internet]. 2021. Available from: http://wsjkw.hebei.gov.cn/xgyqtb/index.jhtml

16. China TCG of the PR of. Guiding Opinions on Scienti c Prevention and Control, Precise Measures, and Divisional Levels for the Prevention and Control of the COVID-19 [Internet]. 2020. Available from: http://www.gov.cn/zhengce/2020-02/18/content_5480514.htm

17. Health commission of Tianjin [Internet]. 2021. Available from: http://wsjk.tj.gov.cn/ZTZL1/ZTZL750/YQFKZL9424/

18. Provisions on the statistical division of urban and rural areas in China [Internet]. 2020. Available from: http://www.stats.gov.cn/tjsj/pcsj/rkpc/5rp/html/append7.htm

19. Blumberg S, Lloyd-Smith JO. Inference of R0 and Transmission Heterogeneity from the Size Distribution of Stuttering Chains. PLoS Comput Biol [Internet]. Public Library of Science; 2013 [cited 2020 Jul 15];9. Available from: /pmc/articles/PMC3642075/?report=abstract
20. Blumberg S, Funk S, Pulliam JRC. Detecting Differential Transmissibilities That Affect the Size of Self-Limited Outbreaks. PLoS Pathog. 2014;10.

21. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. Ijcai. Montreal, Canada; 1995. p. 1137–45.

22. Lau MSY, Grenfell B, Thomas M, Bryan M, Nelson K, Lopman B. Characterizing superspreading events and age-specific infectiousness of SARS-CoV-2 transmission in Georgia, USA. Proc Natl Acad Sci [Internet]. National Academy of Sciences; 2020;117:22430–5. Available from: https://www.pnas.org/content/117/36/22430

23. Qi S, Zhao X, Hao P, Liu N, Gao GF, Song Y, et al. Two Reemergent Cases of COVID-19 — Hebei Province, China, January 2, 2021. China CDC Wkly [Internet]. 3:25–7. Available from: http://weekly.chinacdc.cn//article/id/8d1e2fdd-6bca-4bdf-a700-9ec4eb4a9ae

24. Group C-19 FR, Group LT, Song Y, Zhao X, Li X, Xu W. A Case of COVID-19 — Tianjin Municipality, China, November 7, 2020. China CDC Wkly [Internet]. 2:884–5. Available from: http://weekly.chinacdc.cn//article/id/5ea36aba-015d-48b4-906b-4bfee4c1f69f

25. Rivera-Hernandez M, Ferdows NB, Kumar A. The Impact of the Covid-19 Epidemic on Older Adults in Rural and Urban Areas in Mexico. Journals Gerontol Ser B. 2020;
26. Hu S, Wang W, Wang Y, Litvinova M, Luo K, Ren L, et al. Infectivity, susceptibility, and risk factors associated with SARS-CoV-2 transmission under intensive contact tracing in Hunan, China. medRxiv. Cold Spring Harbor Laboratory Press; 2020;

27. Sun K, Wang W, Gao L, Wang Y, Luo K, Ren L, et al. Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. Science (80- ). 2021;371.

28. Mahale P, Rothfuss C, Bly S, Kelley M, Bennett S, Huston SL, et al. Multiple COVID-19 Outbreaks Linked to a Wedding Reception in Rural Maine—August 7–September 14, 2020. Morb Mortal Wkly Rep. Centers for Disease Control and Prevention; 2020;69:1686.

29. Yang Q, Xiao X, Gu X, Liang D, Cao T, Mou J, et al. Surveillance of common respiratory infections during COVID-19 pandemic demonstrates preventive effectiveness of non-pharmaceutical interventions. Int J Infect Dis. Elsevier; 2021;

30. Markel H, Lipman HB, Navarro JA, Sloan A, Michalsen JR, Stern AM, et al. Nonpharmaceutical interventions implemented by US cities during the 1918-1919 influenza pandemic. Jama. American Medical Association; 2007;298:644–54.

31. Widge AT, Rouphael NG, Jackson LA, Anderson EJ, Roberts PC, Makhene M, et al. Durability of responses after SARS-CoV-2 mRNA-1273 vaccination. N Engl J Med. Mass Medical Soc; 2021;384:80–2.
SARS-CoV-2 transmission chains. Reconstructed transmission chains of 942 SARS-CoV-2-infected individuals in the rural outbreak in Hebei province. Each node in the network represents a patient infected with SARS-CoV-2, and each link represents an infector–infectee relationship. The color of the node
denotes the reporting contact type of the infected individuals. The size of the node corresponds to the number of secondary cases.

Figure 2

Comparison of NPIs and vaccination in SARS-CoV-2 transmission. (a) Impact of vaccination (under different levels of the vaccinated population: 20%, 55%, 65%, and 75%) (b) Comparison of vaccination and non-pharmaceutical interventions

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterial.docx