Transmission of COVID-19 from community to healthcare agencies and back to community: a retrospective study of data from Wuhan, China

Mei Yang, Anshu Li, Gengchen Xie, Yanhui Pang, Xiaoci Zhou, Qiman Jin, Juan Dai, Yaqiong Yan, Yan Guo, Xinghua Liu

ABSTRACT

Background The early spatiotemporal transmission of COVID-19 remains unclear. The community to healthcare agencies and back to community (CHC) model was tested in our study to simulate the early phase of COVID-19 transmission in Wuhan, China.

Methods We conducted a retrospective study. COVID-19 case series reported to the Municipal Notifiable Disease Report System of Wuhan from December 2019 to March 2020 from 17 communities were collected. Cases from healthcare workers (HW) and from community members (CM) were distinguished by documented occupations. Overall spatial and temporal relationships between HW and CM COVID-19 cases were visualised. The CHC model was then simulated. The turning point separating phase 1 and phase 2 was determined using a quadratic model. For phases 1 and 2, linear regression was used to quantify the relationship between HW and CM COVID-19 cases.

Results The spatial and temporal distributions of COVID-19 cases between HWS and CMs were closely correlated. The turning point was 36.85±18.37 (range 15–70). The linear model fitted well for phase 1 (mean R²=0.98) and phase 2 (mean R²=0.93). In phase 1, the estimated βs were positive (from 18.03 to 94.99), with smaller βs (from 2.98 to 15.14); in phase 2, the estimated βs were negative (from –4.22 to –81.87), with larger βs (from 5.37 to 78.12).

Conclusion Transmission of COVID-19 from the community to healthcare agencies and back to the community was confirmed in Wuhan. Prevention and control measures for COVID-19 in hospitals and among HWS are crucial and warrant further attention.

INTRODUCTION

COVID-19, caused by SARS-CoV-2, has changed people’s life worldwide. Available evidence suggests that the virus can spread from person to person by droplets, fine aerosols and fomites, as well as with close personal contact in familial, social and external environments. Moreover, the virus does transmit randomly or evenly through penetrate barriers following various venues in favour of transmission to a whole population. Many studies have been conducted to simulate the spatiotemporal transmission dynamics of COVID-19. A study from Ohio, USA showed that there is substantial spatial variation in the spread of the disease, with localised areas showing marked differences in disease attack rates. In addition, results from a study in South Asia revealed spatial hot spot along with a descriptive output on different parameters of COVID-19. However, these studies cannot reflect the characteristics of individuals. Using individual patients’ trajectory data, Cheng et al. revealed that there is a strong positive correlation between transmission of COVID-19 and mobility of the general public. However, in depth, based on daily surveillance and antiepidemic actions in Wuhan, we find that a patient’s occupation is very important in COVID-19 transmission. Namely, in the early period, an important two-phase COVID-19 transmission model—transmission from communities to healthcare agencies and back to communities (CHC)—might exist.
The CHC model proposes that when infected members from the community visit a healthcare agency during the early period, they become a source of infection for healthcare workers (HW). Later, when infected HWs get home after work, they bring the virus back to the community, infecting other community members (CM). According to the CHC model, based on observations, the COVID-19 epidemic among CMs and HWs forms two distinctive phases: phase 1 and phase 2 (Figure 1). In phase 1, there is faster spread of infections among HWs due to very short distances and close contacts between HWs and infected CMs. On the contrary, during the same period, there is slower spread of infections among CMs. In phase 2, the virus spreads slower among HWs but faster among CMs. With increased awareness of COVID-19, healthcare agencies and HWs are often the first to adopt evidence-based preventive measures quickly and strictly, which can significantly slow down the speed of virus transmission. However, it takes more time and requires more efforts for all CMs to adopt any preventive behaviours. Therefore, it can be inferred that even with the same number of infected, the epidemic will grow quicker among CMs than among HWs. A meta-analysis of 11 studies showed that 10.1% of COVID-19-positive patients are HWs. However, the spatial and temporal distributions among HWs and CMs remain unclear.

Although the CHC model is commonly used in communicable diseases, reported study has merely empirically

Figure 1  A proposed transmission model from the community to healthcare agency and back to the community for the novel pathogen SARS-CoV-2.

Table 1  Participant characteristics

| Variable                      | Healthcare workers | Community members | Total   |
|-------------------------------|--------------------|-------------------|---------|
| Total population in Wuhan (1000) | 127.4              | 10765.5           | 10892.9 |
| COVID-19 cases in Wuhan       | 2900               | 47105             | 50005   |
| Attack rate (per 1000)        | 22.76              | 4.38              | 4.59    |
| Sample ratio, n (%)           | 1138 (39.24)       | 12882 (27.35)     | 14020 (28.04) |
| By severity, n (%)            |                    |                   |         |
| Asymptotic and mild           | 513 (45.08)        | 5876 (45.62)      | 6389 (45.57) |
| Moderate                      | 455 (39.98)        | 4603 (35.73)      | 5058 (36.08) |
| Severe                        | 170 (14.94)        | 2403 (18.65)      | 2573 (18.35) |
| By district, n (%)            |                    |                   |         |
| District 1                    | 88 (7.73)          | 1013 (7.86)       | 1101 (7.85) |
| District 2                    | 177 (15.55)        | 2078 (16.13)      | 2255 (16.08) |
| District 3                    | 214 (18.80)        | 3000 (23.29)      | 3214 (22.92) |
| District 4                    | 115 (10.11)        | 1348 (10.46)      | 1463 (10.44) |
| District 5                    | 442 (38.84)        | 3851 (29.89)      | 4293 (30.62) |
| District 6                    | 102 (8.96)         | 1592 (12.36)      | 1694 (12.08) |
| By sex, n (%)                 |                    |                   |         |
| Male                          | 380 (33.39)        | 5994 (46.53)      | 6374 (45.46) |
| Female                        | 758 (66.61)        | 6888 (53.47)      | 7646 (54.54) |
| By age, n (%)                 |                    |                   |         |
| <35                           | 341 (29.96)        | 1540 (11.96)      | 1881 (13.42) |
| 35–49                         | 443 (38.93)        | 2963 (23.00)      | 3406 (24.29) |
| >50                           | 354 (31.10)        | 8379 (65.04)      | 8733 (62.29) |
| Mean (SD)                     | 39.60 (12.12)      | 55.22 (15.71)     | 53.96 (16.49) |
| Days, symptom to diagnosis    |                    |                   |         |
| Median (IQR)                  | 9 (5–16)           | 10 (5–16)         | 10 (5–16) |
| Mean (SD)                     | 11.52 (8.69)       | 12.50 (8.31)      | 12.43 (8.34) |
tested this hypothesis in COVID-19. If the CHC model is tenable, effective early interventions to control HW to CW transmission would effectively reduce the community transmission of COVID-19. Hence, the purpose of this study is to test the proposed CHC model in COVID-19 to expand evidence on COVID-19 transmission.

**MATERIALS AND METHODS**

The first suspected case of COVID-19 was identified in Wuhan on 8 December 2019.13 A national outbreak was declared on 20 January 2020. Wuhan is the provincial capital of Hubei, with 13 administrative districts and 270 communities, with a total population of 11 million and a total gross domestic product (GDP) of $231,830 million.

**Source of data**

A retrospective study was conducted. Case series of resident patients with COVID-19 reported to the Wuhan Center for Disease Prevention and Control from December 2019 to 26 March 2020 were included. For individual cases, data on occupation and community where they live were also extracted. All data were derived from the Municipal Notifiable Disease Report System of Wuhan (https://wwwcdc.gov/coronavirus/2019-ncov/
the ratio of the total COVID-19 cases to the population at risk. The total number of persons employed by preventive and clinical care agencies, including hospitals, outpatient clinics, emergency centres, pharmacies, and disease control and prevention institution, were classified as HW infections. To better address the study question, non-professional workers in healthcare settings were also considered as HWs, such as security guards, technicians and logistics personnel. Apart from this, other cases were treated as CW infections.

COVID-19 diagnosis and grouping

All COVID-19 cases included in this analysis were confirmed by RT-PCR assay test in the laboratory with a positive SARS-CoV-2. The test was conducted following national standards. Clinic type was categorised as mild, moderate, severe or critical according to the diagnosis and treatment scheme for COVID-19 released by the National Health Commission of China. According to occupation, subjects with COVID-19 were divided into HW and CW infections. If cases were employed by preventive and clinical care agencies, including hospitals, outpatient clinics, emergency centres, pharmacies, and disease control and prevention institution, they were classified as HW infections. To better address the study question, non-professional workers in healthcare settings were also considered as HWs, such as security guards, technicians and logistics personnel. Apart from this, other cases were treated as CW infections.

Statistical analysis

Descriptive analysis was conducted to assess the distribution of cases by age, sex, date of disease onset, clinical severity, location by district and days from first symptom onset to diagnosis. Self-reported date of fever or cough onset was used as the date of disease onset and was applied to construct the onset curves. Attack rate was computed as the ratio of the total COVID-19 cases divided by the total population at risk. The total number of persons employed by healthcare agencies and the total CMs (minus the total HWs) in 2019 were used as the at-risk population for HWs and CMs, respectively. Sample ratio was computed by dividing the total COVID-19 cases included in the analysis over the total number of confirmed cases used in this study.

Geographical mapping was used to describe the geographical distribution of the selected communities. Geographical mapping was conducted using ArcGIS V.10.2 software (Environmental Systems Research Institute). Daily and cumulative COVID-19 cases for HWs and CMs for the 17 communities were first plotted by date to show the overall relationship between the two.

The CHC model was tested by the following process for each community. First, cumulative COVID-19 cases of HW by date were plotted against those of CM to obtain a visual assessment of the two-phase pattern: a slower increase in phase 1 and a quicker increase in phase 2 with a turning point. The trajectories of scatter plots were approximately quadratic curves. Hence, to quantify the turning point, data for individual communities were fitted with the quadratic model:

\[ y = \alpha + \beta_1 x + \beta_2 x^2, \]  
\[ \tau = \frac{\beta_1}{(-2\beta_2)} \]

where \( x \) and \( y \) are daily cumulative COVID-19 cases by date for HW and CM, respectively.

After the turning point \( \tau \) was determined for a community, the data for phase 1 and phase 2 were analysed separately using a linear regression model. Statistical analyses were performed with SPSS Software V.23.0 for Windows.

RESULTS

Demographic and epidemiological characteristics

Overall, the attack rates of COVID-19 per 1000 population were 22.76 (2900/127.4) for HW and 4.38 (47 105/10 765.5) for CM, respectively. Table 1 summarises the characteristics of the study participants from all 17 communities. Of a total of 14020 cases, 91.88% were CM and 8.12% were HW. Compared with HWs, there were more male (46.53% vs 33.39%) and older (≥50 years; 65.04% vs 51.10%) cases among CMs. The median (IQR) duration from showing a symptom to receiving a clinical diagnosis was 9 (5–16) days for HWs and 10 (5–16) days for CMs.

Spatiotemporal relationship between HWs and CMs

Figure 2 depicts the spatiotemporal relationship of COVID-19 between HWs and CMs. As shown in figure 2A, there were close geographical relations between the two groups. The scatter plots in figure 2B (daily new cases) and figure 2C (cumulative cases) show that temporal trends in the COVID-19 epidemic among HWs were exactly similar to those among CMs. In figures 3 and 4, the cumulative cases of HW (right vertical axis) and CM (left vertical axis) were plotted for all 17 communities. All plots showed the two-phase patterns, with slow increase early and a quick increase later.

CHC model simulation

First, the turning points were measured by the quadratic model for each community, as presented in table 2. Overall, the average turning point separating phase 1 from phase 2 was 38.65 (SD=18.37, minimum=15, maximum=70). Linear regression model was then applied to simulate the
relationship between the daily cumulative cases of CM and HW before and after the turning point, namely phase 1 and phase 2. The results showed that the estimated $R^2$ in phase 1 ranged from 0.92 to 0.99 (mean=0.98, SD=0.02), while in phase 2 varied from 0.69 to 0.99 (mean=0.93, SD=0.08). Consistent with the proposed CHC model in figure 1, the estimated $\hat{\alpha}$s in phase 1 were positive for all communities, ranging from 18.03 to 94.99 (mean=47.38, SD=21.41), while the estimated $\alpha$s in phase 2 were all negative, ranging from $-4.22$ to $-81.47$ (mean=$-41.31$, SD=24.49). Moreover, the estimated $\hat{\beta}$s in phase 1 were smaller than those in phase 2 for each community (from 2.98 to 15.14 vs from 5.37 to 78.12). Above all, the results supported the proposed CHC model.

**DISCUSSION**

In this study, we empirically tested the proposed CHC model in the early spread of COVID-19 in Wuhan. The results confirmed the transmission of COVID-19 from the community to healthcare agencies and back to the community. To the best of our knowledge, this study is the first to demonstrate this heterogeneity in the spread of the novel virus.

Different from the theories of enrolling the whole population, including infected and uninfected individuals, to simulate the model, only daily cumulative COVID-19 cases were applied in our CHC model. Furthermore, the core variable is an individual’s occupation. Published literature well analysed the spatiotemporal transmission patterns of COVID-19, however, the influence of individual characteristics has never been evaluated. The findings of our study indicate that the novel virus can be spread through special venues, and effective early interventions to control HW to CW transmission would effectively reduce the community transmission of COVID-19.

Based on the daily surveillance and antiepidemic actions in Wuhan, there is evidence supporting the CHC model. Although an increased risk of COVID-19 infection among medical and health professionals and non-professional workers has been reported, COVID-19 cases are detected earlier among CMs than among HWs. A slower growth of cases is then observed among CMs compared with HWs. Subsequently, there is a clear turning point where CM cases progress faster than HW cases. The progressive knowledge of COVID-19 in HWs and the inadequate community protection in CMs might cause different changes in different populations. Hence, the simulation of the proposed CHC model in COVID-19 is of significance in order to understand its early spatio-temporal transmission pattern.

The CHC model partially explains the surge of COVID-19 during the early period in Wuhan and many other places inside and outside of China. During the
early period of the epidemic, few HWs adopted personal protection equipment (PPE) when taking care of patients and few hospitals imposed protective measures, except specialised hospitals for infectious diseases. Moreover, a lot of doctors experienced difficulties in correctly diagnosing a COVID-19 infection. In quick succession other professionals by sending patients to various departments for further assessment exposed, and then the epidemic among HWs exacerbated.

What is much alarming than the spread among HWs is that there is huge transmission from healthcare agencies back to communities. With lack of knowledge about the epidemic, infected HWs might bring the virus back to the community without taking adequate protection measures. Daily commute between healthcare agencies and communities directly provides an opportunity to unintentionally spread COVID-19.

In addition, the CHC model, of which ‘H’ indicates the core of the spread, partially explains the rapid control of the COVID-19 epidemic in China. With a population of 1.4 billion, China technically controlled the epidemic by 8 March 2020, about 4 months after the first suspected case on 8 December 2019. Soon after the epidemic has been understood, one important measure taken by China was to control the path of transmission from the community to healthcare agencies. Typical measures included (1) strict protection measures among HWs (ie, using PPE, not going home and instead staying at specialised camps after work), (2) prohibition of hospital visits and (3) quarantine of all infected cases in specialised or newly constructed hospitals.

The CHC model also bears significant implications for emerging pathogen epidemics in the future. The greatest threat of a novel epidemic is the transmission of infection among CMs to HWs then back to the CMs due to lack of knowledge about the pathogen during the early stage. The hidden attack of a novel infection makes it hard to take antiepidemic measures for everyone and every agency, particularly for healthcare professionals and agencies, during the early stage of the epidemic.

There are limitations to this study. First, the data used in this study were inadequate for us to link the transmission of the virus at the individual person level to prove the CHC model. Second, this study was a secondary analysis. Data on other factors at the community and individual levels that may affect the two-phase CHC model were not included, such as community type, population size, economic level, knowledge about the epidemic, implementation of prevention measures by community and healthcare agencies, use of PPE, and frequency of contact between the infected and the uninfected within HWs and CMs and between HWs and CMs.

Despite the limitations, this study is the first to test the CHC model in the COVID-19 epidemic, with the
results highlighting that blocking the transmission between healthcare agencies and communities is critical to protecting invaluable HWs and slowing down the epidemic in a population. In addition to advancing our understanding of the COVID-19 epidemic and its control, findings of this study add new data that can guide the control and prevention of a novel pathogenic epidemic and underscore the need for further research to explore the underlying dynamics of the COVID-19 epidemic.

**CONCLUSIONS**

In summary, transmission of COVID-19 from the community to healthcare and back to the community was confirmed in Wuhan.

**Author affiliations**

1. Department of Maternal and Child Health, Wuhan University of Science and Technology, Wuhan, China
2. Department of Gastrointestinal Surgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
3. Department of Information Center, Wuhan Center for Disease Control and Prevention, Wuhan, China
4. Chronic Noncommunicable Diseases Control and Prevention, Wuhan Center for Disease Control and Prevention, Wuhan, China

**Acknowledgements** The authors gratefully acknowledge the COVID-19 data reporting and administrative staff.

**Contributors** MY, AL and GX analysed the data and finalised the manuscript. YG and XL designed the project, analysed the data and prepared the first draft. YP prepared the figures and tables. XZ, QJ and JD collected the information. YY designed the project and collected the information. XL is responsible for the overall content as the guarantor.

**Funding** The study was funded by the Wuhan Municipal Health Commission (grant number: EG20A04) and the Hubei Municipal Health Commission (grant number: WJ2019F011).

**Disclaimer** The funders had no role in the design and conduct of the study; collection, management, analysis and interpretation of data; preparation, review or approval of the manuscript; and decision to submit the manuscript for publication.

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not required.

---

**Table 2** Relationship of cumulative daily COVID-19 cases between community members and healthcare workers in phase 1 and phase 2

| Community | Turning point | Before |  | After |  |
|-----------|---------------|--------|---|-------|---|
|           |  | $\hat{\alpha}$ | $\hat{\beta}$ | $R^2$ | $\hat{\alpha}$ | $\hat{\beta}$ | $R^2$ |
| 1         | 62            | 72.71  | 7.92 | 0.99 | -68.56 | 18.02 | 0.99 |
| 2         | 60            | 77.66  | 9.85 | 0.99 | -81.87 | 16.03 | 0.99 |
| 3         | 43            | 39.69  | 7.50 | 0.97 | -40.92 | 9.40  | 0.97 |
| 4         | 70            | 35.66  | 5.29 | 0.98 | -10.56 | 78.12 | 0.69 |
| 5         | 40            | 36.65  | 15.14| 0.97 | -22.97 | 16.29 | 0.91 |
| 6         | 23            | 32.05  | 5.16 | 0.97 | -4.22  | 23.51 | 0.97 |
| 7         | 70            | 68.12  | 10.30| 0.99 | -65.47 | 11.94 | 0.99 |
| 8         | 42            | 35.47  | 5.66 | 0.98 | -42.77 | 11.18 | 0.97 |
| 9         | 32            | 49.22  | 10.26| 0.98 | -62.87 | 17.84 | 0.99 |
| 10        | 50            | 94.99  | 2.98 | 0.99 | -50.51 | 9.81  | 0.98 |
| 11        | 38            | 44.20  | 3.01 | 0.98 | -16.29 | 5.37  | 0.82 |
| 12        | 21            | 25.61  | 6.47 | 0.96 | -26.53 | 12.44 | 0.92 |
| 13        | 24            | 65.88  | 9.13 | 0.99 | -73.35 | 11.01 | 0.99 |
| 14        | 23            | 39.79  | 4.46 | 0.99 | -28.15 | 9.71  | 0.93 |
| 15        | 15            | 21.99  | 11.89| 0.95 | -29.25 | 15.80 | 0.93 |
| 16        | 29            | 47.79  | 10.08| 0.99 | -14.37 | 17.81 | 0.80 |
| 17        | 15            | 18.03  | 7.48 | 0.92 | -63.65 | 17.27 | 0.98 |
| Range     | 15, 70        | 18.03,94.99 | 2.98, 15.14 | 0.92, 0.99 | -4.22, -81.87 | 5.37, 78.12 | 0.69, 0.99 |
| Mean      | 38.65         | 47.38  | 7.80 | 0.98 | -41.31 | 17.74 | 0.93 |
| SD        | 18.37         | 21.41  | 3.27 | 0.02 | 24.49  | 16.18 | 0.08 |

The estimated intercept and slope were all statistically significant at $p<0.01$ and $R^2$ was for assessment of data model fit. The turning point was the number of cumulative COVID-19 cases among healthcare workers (see text for details).

Following the proposed CHC model, the $\hat{\alpha}$ was positive before the turning point and negative after the point; and the $\hat{\beta}$ before the point was smaller than after the point (see figure 1). The difference in the mean $\hat{\alpha}$ and $\hat{\beta}$ before and after the turning point was statistically significant ($p<0.01$ for both).

CHC, community to healthcare agencies and back to community.
Ethics approval This study was approved by the Ethics Committee of the Wuhan Center for Disease Prevention and Control (approval number: WHCDCIRB-K-2020017). The results were reported according to the STROBE statement. All methods were performed in accordance with the relevant guidelines and regulations. Epidemiological data were obtained from the Municipal Notifiable Disease Report System and did not involve personal information. The Ethics Committee of the Wuhan Center for Disease Prevention and Control had waived the requirement for informed consent for this study.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD
Xinghua Liu http://orcid.org/0000-0001-8783-0573

REFERENCES
1 Centers for Disease Control and Prevention. How COVID-19 spreads. Available: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/how-covid-spreads.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fpreparing%2Ftransmission.html
2 Li W, Zhang B, Lu J. The characteristics of household transmission of COVID-19. Clin Infect Dis 2020.
3 Lu C-W, Liu X-F, Jia Z-F. 2019-nCoV transmission through the ocular surface must not be ignored. Lancet 2020;395:e39.
4 Willcox MD, Walsh K, Nichols JJ, et al. The ocular surface, coronaviruses and COVID-19. Clin Exp Optom 2020;103:148–24.
5 Yeo C, Kaushal S, Yeow D. Enteric involvement of coronaviruses: is faecal-oral transmission of SARS-CoV-2 possible? Lancet Gastroenterol Hepatol 2020;5:335–7.
6 Chu DK, Akil EA, Duda S, et al. 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;395:1973–87.
7 Shen Y, Xu W, Li C, et al. A cluster of novel coronavirus disease 2019 infections indicating person-to-person transmission among casual contacts from social gatherings: an outbreak Case-Contact investigation. Open Forum Infect Dis 2020;7:ofi231.
8 Cuadros DF, Xiao Y, Mukandavire Z, et al. Spatiotemporal transmission dynamics of the COVID-19 pandemic and its impact on critical healthcare capacity. Health Place 2020;64:102404.
9 Islam A, Sayeed MA, Rahman MK, et al. Spatiotemporal patterns and trends of community transmission of the pandemic COVID-19 in South Asia. Bangladesh as a case study. Biosaf Health 2021;1:39–49.
10 Cheng T, Lu T, Liu Y, et al. Revealing spatiotemporal transmission patterns and stages of COVID-19 in China using individual patients’ trajectory data. Comput Urban Sci. 2021:1.