

Household secondary attack rate of COVID-19 and associated determinants in Guangzhou, China: a retrospective cohort study



Qin-Long Jing*, Ming-Jin Liu*, Zhou-Bin Zhang*, Li-Qun Fang*, Jun Yuan*, An-Ran Zhang, Natalie E Dean, Lei Luo, Meng-Meng Ma, Ira Longini, Eben Kenah, Ying Lu, Yu Ma, Neda Jalali, Zhi-Cong Yang, Yang Yang

Summary

Background As of June 8, 2020, the global reported number of COVID-19 cases had reached more than 7 million with over 400 000 deaths. The household transmissibility of the causative pathogen, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), remains unclear. We aimed to estimate the secondary attack rate of SARS-CoV-2 among household and non-household close contacts in Guangzhou, China, using a statistical transmission model.

Methods In this retrospective cohort study, we used a comprehensive contact tracing dataset from the Guangzhou Center for Disease Control and Prevention to estimate the secondary attack rate of COVID-19 (defined as the probability that an infected individual will transmit the disease to a susceptible individual) among household and non-household contacts, using a statistical transmission model. We considered two alternative definitions of household contacts in the analysis: individuals who were either family members or close relatives, such as parents and parents-in-law, regardless of residential address, and individuals living at the same address regardless of relationship. We assessed the demographic determinants of transmissibility and the infectivity of COVID-19 cases during their incubation period.

Findings Between Jan 7, 2020, and Feb 18, 2020, we traced 195 unrelated close contact groups (215 primary cases, 134 secondary or tertiary cases, and 1964 uninfected close contacts). By identifying households from these groups, assuming a mean incubation period of 5 days, a maximum infectious period of 13 days, and no case isolation, the estimated secondary attack rate among household contacts was 12.4% (95% CI 9.8–15.4) when household contacts were defined on the basis of close relatives and 17.1% (13.3–21.8) when household contacts were defined on the basis of residential address. Compared with the oldest age group (≥ 60 years), the risk of household infection was lower in the youngest age group (< 20 years; odds ratio [OR] 0.23 [95% CI 0.11–0.46]) and among adults aged 20–59 years (OR 0.64 [95% CI 0.43–0.97]). Our results suggest greater infectivity during the incubation period than during the symptomatic period, although differences were not statistically significant (OR 0.61 [95% CI 0.27–1.38]). The estimated local reproductive number (R) based on observed contact frequencies of primary cases was 0.5 (95% CI 0.41–0.62) in Guangzhou. The projected local R , had there been no isolation of cases or quarantine of their contacts, was 0.6 (95% CI 0.49–0.74) when household was defined on the basis of close relatives.

Interpretation SARS-CoV-2 is more transmissible in households than SARS-CoV and Middle East respiratory syndrome coronavirus. Older individuals (aged ≥ 60 years) are the most susceptible to household transmission of SARS-CoV-2. In addition to case finding and isolation, timely tracing and quarantine of close contacts should be implemented to prevent onward transmission during the viral incubation period.

Funding US National Institutes of Health, Science and Technology Plan Project of Guangzhou, Project for Key Medicine Discipline Construction of Guangzhou Municipality, Key Research and Development Program of China.

Copyright © 2020 Elsevier Ltd. All rights reserved.

Introduction

The ongoing COVID-19 pandemic, caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has now affected 188 countries worldwide. As of June 8, 2020, more than 7 million reported cases and over 400 000 deaths had been reported.¹ Older individuals (aged ≥ 70 years) and individuals with chronic conditions such as diabetes and cardiopulmonary disease are most susceptible to severe disease and death.² Efficient viral transmission via droplets and fomites is potentially

supplemented by other transmission routes such as aerosol and faecal contamination.^{3,4} Accumulating evidence suggests that presymptomatic or asymptomatic carriers can transmit the virus.^{5–7} Within-household transmission is suspected to have contributed substantially to the continued increase in cases in China following the introduction of nationally enforced restrictions on human movement.^{8,9} Isolation of cases and quarantine of their close contacts at home are frequently recommended as a disease control measure in countries with COVID-19 outbreaks, but such

Lancet Infect Dis 2020; 20: 1141–50

Published Online
June 17, 2020
[https://doi.org/10.1016/S1473-3099\(20\)30471-0](https://doi.org/10.1016/S1473-3099(20)30471-0)

See [Comment](#) page 1103

For the Chinese translation of the abstract see [Online](#) for appendix 1

*Contributed equally

Guangzhou Center for Disease Control and Prevention, Guangzhou, China

(Q-L Jing PhD, Z-B Zhang MPH, J Yuan MD, L Luo PhD,

M-M Ma MD, Y Lu MD, Y Ma MD,

Z-C Yang MD); Department of

Biostatistics, College of Public

Health and Health Professions,

Emerging Pathogens Institute,

University of Florida,

Gainesville, FL, USA (M-J Liu BS,

A-R Zhang BS, N E Dean BS,

I Longini PhD, N Jalali MS,

Y Yang PhD); State Key

Laboratory of Pathogen and

Biosecurity, Beijing Institute of

Microbiology and

Epidemiology, Beijing, China

(A-R Zhang, L-Q Fang PhD);

Department of Epidemiology,

School of Public Health,

Cheeloo College of Medicine,

Shandong University, Jinan,

China (A-R Zhang); and Division

of Biostatistics, College of

Public Health, Ohio State

University, Columbus, OH, USA

(E Kenah ScD)

Correspondence to:

Dr Yang Yang, Department of

Biostatistics, College of Public

Health and Health Professions,

Emerging Pathogens Institute,

University of Florida, Gainesville,

FL 32611, USA

yangyang@ufl.edu

or

Dr Zhi-Cong Yang, Guangzhou

Center for Disease Control and

Prevention, Guangzhou 510440,

China

yangzcgzcdc.org.cn

Research in context**Evidence before this study**

The transmissibility of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in household and community settings remains underinvestigated. On April 1, 2020, we searched PubMed and medRxiv using the search terms ("novel coronavirus" OR "COVID-19" OR "SARS-CoV-2" OR) AND ("household" OR "family") AND ("transmissibility" OR "attack rate"). Our search yielded three articles that investigated household secondary attack rates of COVID-19 in multiple family clusters. In studies of contact tracing data in Shenzhen and Guangzhou, two cities in southern China, the secondary attack rates were 14.9% and 10.2%, respectively, but these estimates represented the proportions of infections among household contacts. In a study of the close contacts of ten US patients with COVID-19 the estimated household secondary attack rate was 10.5%, however, the sample size was too small for reliable interpretation and only symptom onset of primary cases was examined. Transmission events from COVID-19 cases during their incubation period or from asymptomatic carriers have been reported, but infectiousness before symptom onset has not been quantified.

Added value of this study

Using contact tracing data from patients with COVID-19 in Guangzhou, China, we implemented a statistical transmission model to estimate secondary attack rate for household and non-household close contacts. To our knowledge, our model is the first to account for heterogeneous individual-level exposure history, tertiary transmission, potential exposure to untraced infection sources, and asymptomatic infections. Additionally, we assessed the effects of age, sex, epidemic phase, and household size on transmissibility of the virus and relative infectivity before and after symptom onset.

Implications of all the available evidence

SARS-CoV-2 can be transmitted efficiently within households and during the incubation period of COVID-19 cases. Because presymptomatic and asymptomatic transmission has been observed, case-isolation alone is inadequate for mitigating the pandemic. Comprehensive tracing and timely quarantine of close contacts of COVID-19 cases should be implemented to prevent onward transmission during their incubation periods.

restrictions are likely to have little or no effect on transmission within households.

To date, transmissibility of the disease has primarily been assessed at the population level, using mathematical models, or at the individual level in synthetic populations using agent-based models coupled with statistical methods.¹⁰⁻¹² Transmissibility within households or through other types of close contact remains underinvestigated, despite the importance of these social interactions in shaping the overall dynamics of disease spread and in determining the effectiveness of mitigation strategies.¹³ Data obtained from contact tracing provide the most accurate information about human-to-human transmissibility of any infectious pathogen, because transmissibility can be assessed more accurately by accounting for individual-level exposure history. Available estimates for the secondary attack rate (defined as the probability that an infected individual will transmit the disease to a susceptible individual [eg, household or close contacts]) of SARS-CoV-2 are based on contact tracing data for hundreds of cases in Shenzhen and Guangzhou in China, and ten cases in the USA.¹⁴⁻¹⁶ These estimates represent the proportion of confirmed infections among all traced contacts, which does not fully account for heterogeneity in individual exposure history, the possibility of transmission among contacts themselves, or the infection risks from untraced contacts or fomites.

We aimed to estimate the secondary attack rate of SARS-CoV-2 among household and non-household close contacts in Guangzhou, using a statistical transmission model. This model accounts for individual-level exposure history and the potential existence of tertiary cases. We also aimed to assess the effects of age and sex on the infectivity

of COVID-19 cases and susceptibility of their close contacts, and the relative infectivity of COVID-19 cases during the incubation period compared with the period of illness.

Methods**Case definition**

A suspected COVID-19 case was defined as an individual who met one or more epidemiological criteria (had travelled to, or resided in, Wuhan or nearby cities in the 14 days before symptom onset; had contact history with an individual with COVID-19 [confirmed by RT-PCR] in the 14 days before symptom onset; had contact history with patients who had fever or respiratory symptoms and came from Wuhan or communities with reported COVID-19 cases in the 14 days before symptom onset; or was related to a cluster of COVID-19 cases) and two or more clinical criteria (had fever or respiratory symptoms; had radiographical characteristics of pneumonia; or normal or reduced leucocyte counts, or reduced lymphocyte counts during the acute phase of the disease; appendix 2 p 1). A confirmed case was defined as a suspected case with positive detection of SARS-CoV-2 nucleic acid by real-time RT-PCR or viral genes that are highly homologous to SARS-CoV-2 by sequencing using respiratory specimens. An individual with laboratory confirmation, but without clinical signs, mainly found by community screening and contact tracing, was considered asymptomatic (appendix 2 pp 2, 3). In this study, asymptomatic infections were analysed as confirmed cases.

Epidemiological investigation and contact tracing

Epidemiological investigations were done by county-level Centers for Disease Control and Prevention (CDC)

See Online for appendix 2

offices in China within 24 h after a suspected or confirmed case was reported (appendix 2 pp 2,3). For each suspected or confirmed case, we recorded demographic, clinical, diagnostic, and occupational data, baseline health conditions, clinical samples and laboratory test results, and exposure history in the 14 days before symptom onset using a standardised investigation form. Typically, a suspected case would be changed to a confirmed case or removed from the national surveillance system once laboratory test results were available. A close contact was defined as an individual who had unprotected close contact (within 1 m) with a confirmed case within 2 days before their symptom onset or sample collection, including but not limited to household members, caregivers, and individuals in the same workplace, classroom, hospital ward, or transportation vehicle. Close contacts were quarantined at designated places (eg, hotel rooms) or at home and followed up for 14 days, and nasal swabs were collected at day 1 and day 14 and tested by RT-PCR (appendix 2 pp 3, 4). These data were collected as part of a continuing public health response required by the National Health Commission of China, thus the requirement for written informed consent was waived. This study was approved by the ethics committee of the Guangzhou CDC (Guangzhou, China). Analyses involving personally identifiable data were done at Guangzhou CDC. Anonymised data were used for all other analyses.

In our analyses, individuals who were linked by contact tracing were considered a close contact group. Cases in the same close contact group formed a case cluster. An imported case was defined as an individual who resided in or had travelled to Hubei province (of which Wuhan is the capital) in the 2 weeks before symptom onset; otherwise, individuals were considered a local case. In a close contact group, a local case with symptom onset 1 day or less (3 days or less for an imported case) from the earliest onset day in the close contact group was considered a primary case; otherwise, this case was considered a secondary case. For asymptomatic infections, the primary or secondary case status was determined by the collection dates of the earliest SARS-CoV-2 positive specimens. We used two definitions of household contacts: individuals who were either family members or close relatives, such as parents and parents-in-law, regardless of residential address, and individuals living at the same address regardless of relationship.

Statistical analysis

We used standard non-parametric tests (Fisher's exact test, the Kruskal-Wallis test, and the Mann-Whitney U test) to compare characteristics between demographic groups. Similar to the published literature,¹⁴⁻¹⁶ we calculated the proportions of confirmed infections among all traced contacts for subgroups and herein, we refer to these proportions as data-based secondary attack rate estimates, which do not account for the fact that

infections among contacts are not necessarily secondary, and could be tertiary. All summary analyses were done using R statistical software (version 3.6.1). The spatial distribution of case clusters was mapped at the community level using ArcGIS (Environmental Systems Research Institute, Redlands, CA, USA), with a directed graph indicating potential transmission chains. We estimated effective reproductive numbers (R_e) based on the contact tracing data (appendix 2 pp 4-6). Since the transmission relationship remains unclear, we investigated three scenarios defined by the following assumptions: scenario 1, all imported cases were primary cases, and all secondary cases were infected by primary cases in the same case cluster; scenario 2, which was identical to scenario 1, with the additional assumption that local primary cases might have been infected by earlier cases in other case clusters; and scenario 3, which was identical to scenario 2, with the additional assumption that imported secondary cases were considered secondary cases rather than primary cases. Scenarios 1 and 3 served as the lower and upper bounds of the R_e .

A chain-binomial statistical model was used to estimate secondary attack rate and local reproductive number, with an expectation-maximisation algorithm to account for uncertainty in the infection time of asymptomatic infections (appendix 2 pp 6-10).¹⁷ Possible distributions of the incubation period and the infectious period were derived from the literature and our previous research on a separate contact tracing database (appendix 2 pp 10-12).^{5,18-20} On the basis of the available literature (appendix 2 p 10-11), we reported the estimates associated with a mean incubation period of 5 days and a maximum infectious period of 13 days as the primary results.

Briefly, this model estimated the probabilities, p_1 and p_2 , of viral transmission from an infectious household contact and from an infectious non-household contact (eg, friends, coworkers, passengers) respectively, to a susceptible person per daily contact. Additionally, each susceptible person was subject to a constant daily probability, b , of being infected by an unspecified external source, which accounted for untraced contacts and fomites. We assumed that the infectivity of a COVID-19 case differs between the incubation period and the period after symptom onset (ie, illness period). We used (D_{\min} , D_{\max}) to represent the whole infectious period with the symptom onset day set as day 0, and modelled the effective daily transmission probability as

$$p_k^*(l) = p_k$$

on days $D_{\min} \leq l < 0$ (incubation period) and

$$\frac{p_k^*(l)}{1 - p_k^*(l)} = \frac{p_k}{1 - p_k} \times \text{OR}$$

on days $0 \leq l < D_{\max}$ (illness period), where the odds ratio (OR) measures the relative infectivity of the illness period versus the incubation period. The secondary attack rate was defined as:

$$1 - \prod_{l=D_{\min}}^{D_{\max}} [1 - p_k^*(l)\phi^*(l)], \quad k=1,2,$$

where $\phi^*(l)$ was the prespecified relative infectivity level on the l^{th} day of the infectious period based on previous studies,¹⁹ peaking near the time of symptom onset. Herein, we refer to the derived secondary attack rate estimates as model-based. The local reproductive number (ie, the mean number of infections a symptomatic case could generate via household and non-household contacts) was defined as:

$$\sum_{k=1}^2 \sum_{l=D_{\min}}^{D_{\max}} \{n_k(l)p_k^*(l)\phi^*(l)\prod_{m=D_{\min}}^{l-1} [1 - p_k^*(l)\phi^*(l)]\}$$

where $n_1(l)$ and $n_2(l)$ were the mean numbers of household and non-household contacts per primary case on day l (appendix 2 pp 12, 13).

We assessed the effect of age, sex, household size, and epidemic phase on susceptibility and infectivity by regressing transmission probabilities on these characteristics associated with either the susceptible person or the infectious person in each potential transmission–exposure pair (appendix 2 pp 13, 14). Additionally, we assessed the goodness-of-fit of the model under different settings of the natural history of disease (appendix 2 p 14,15).

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or the writing of the report. The corresponding authors had full access to all the data in the study, and had final responsibility for the decision to submit for publication.

Results

Between Jan 7, 2020, and Feb 18, 2020, 349 laboratory-confirmed SARS-CoV-2 infections were reported to Guangzhou CDC, among whom 19 (5%) individuals were asymptomatic. Contact tracing identified 195 unrelated close contact groups (215 primary cases, 134 secondary or tertiary cases, and 1964 uninfected close contacts). The median size of close contact groups was six (IQR 4–10; range 1–274). For 138 (71%) of these 195 close contact groups, no secondary cases were identified. Time from symptom onset to hospital admission and to laboratory confirmation was longer for primary cases than for secondary cases and was longer in January than in February (appendix 2 p 19). Among the 349 cases, the most common symptoms were fever (258 [74%] cases), cough (207 [59%]), fatigue (74 [21%]), sore throat (60 [17%]), chills (55 [16%]), and myalgia (51 [15%]). Radiographic abnormality was observed in 238 (80%) of 349 cases (appendix 2 p 20).

Most patients with COVID-19 were adults (aged 20–59 years; table 1). The majority of primary cases (158 [73%] of 215 cases) and nearly half of all secondary cases (66 [46%] of 134 cases) had recently travelled to or resided in Hubei province (referred to as imported cases hereafter). The overall data-based secondary attack rates were 13.2% (95% CI 10.9–15.7) among household contacts and 2.4% (1.6–3.3) among non-household contacts, when household was defined on the basis of close relatives. Within households, the data-based secondary attack rates were lower in the youngest age group (age <20 years; 5.2% [95% CI 2.4–9.7]) than the 20–59 years age group (14.8% [95% CI 11.7–18.4]; $p=0.0009$) and the oldest age group (age ≥ 60 years; 18.4% [12.5–25.6%]; $p=0.0003$). Similar findings were observed among non-household contacts, but the differences between age groups were not statistically significant. We identified no significant differences in secondary attack rate between sexes. Secondary attack rate estimates decreased between January and February, 2020, within and outside households (both $p < 0.0001$). When household was defined by residential address, the data-based secondary attack rate among household contacts increased to 17.2% (95% CI 14.1–20.6) and was 2.6% (1.9–3.6) among non-household contacts.

Most COVID-19 cases were reported in the densely populated districts (where 56% of the total population of Guangzhou reside) including Yuexiu, Liwan, Haizhu, Tianhe, and Baiyun (figure 1). Four clusters with five or more secondary cases (excluding tertiary cases or further generations) were identified, with one cluster identified in Yuexiu, Haizhu, Baiyun and Panyu districts, and all primary cases in these clusters were imported (figure 1B–D). The longest transmission chain had three generations subsequent to the primary case, which occurred in the Panyu district (figure 1D). Five other clusters had two subsequent generations (figure 1B, figure 1C). Most of the reported residential locations of primary and secondary cases within the same clusters were identical, but some non-household secondary cases resided in areas that were distant from the primary cases. Most transmissions occurred between household members (appendix 2 p 28).

The first imported primary case had symptom onset on Jan 7, 2020, and arrived in Guangzhou on Jan 13, 2020. The earliest local primary case had symptom onset on Jan 16, 2020, when at least two imported cases had already been reported in Guangzhou. The number of imported cases peaked at the same time as the epidemic in Guangzhou around Jan 27, 2020, 4 days after the lockdown was implemented in Wuhan (figure 2). After Jan 27, 2020, the number of imported cases decreased and the epidemic waned quickly, with only sporadic cases reported by the middle of February. In the early phase of the epidemic, the R_t reached 1.4 in scenario 3, around 1.2 in scenario 2, and 0.7 in scenario 1. For all scenarios, by Jan 27, 2020, the R_t had declined to less than 0.5,

	Primary cases	Secondary cases		Uninfected close contacts		Overall	Data-based secondary attack rate*	
		Household	Non-household	Household	Non-household		Household	Non-household
Close relatives								
Age, years								
<20	10/215 (5%)	9/103 (9%)	1/31 (3%)	163/681 (24%)	70/1283 (5%)	253/2313 (11%)	5.2% (2.4–9.7)	1.4% (0.04–7.6)
20–59	145/215 (67%)	67/103 (65%)	22/31 (71%)	385/681 (57%)	961/1283 (75%)	1580/2313 (68%)	14.8% (11.7–18.4)	2.2% (1.4–3.4)
≥60	60/215 (28%)	27/103 (26%)	8/31 (26%)	120/681 (18%)	247/1283 (19%)	462/2313 (20%)	18.4% (12.5–25.6)	3.1% (1.4–6.1)
Sex								
Female	107/215 (50%)	57/103 (55%)	17/31 (55%)	341/681 (50%)	627/1283 (49%)	1149/2313 (50%)	14.3% (11.0–18.2)	2.6% (1.5–4.2)
Male	108/215 (50%)	46/103 (45%)	14/31 (45%)	335/681 (49%)	651/1283 (51%)	1154/2313 (50%)	12.1% (9.0–15.8)	2.1% (1.2–3.5)
Month†								
January	193/215 (90%)	98/103 (95%)	29/31 (94%)	545/681 (80%)	681/1283 (53%)	1546/2313 (67%)	15.2% (12.6–18.3)	4.1% (2.8–5.8)
February	22/215 (10%)	5/103 (5%)	2/31 (6%)	136/681 (20%)	602/1283 (47%)	767/2313 (33%)	3.5% (1.2–8.1)	0.33% (0.04–1.2)
Household size								
≤6 people	160/215 (74%)	69/103 (67%)	23/31 (74%)	302/681 (44%)	874/1283 (68%)	1428/2313 (62%)	18.6% (14.8–22.9)	2.6% (1.6–3.8)
>6 people	55/215 (26%)	34/103 (33%)	8/31 (26%)	379/681 (56%)	409/1283 (32%)	885/2313 (38%)	8.2% (5.8–11.3)	1.9% (0.8–3.7)
Origin								
Imported	158/215 (73%)	59/103 (57%)	3/31 (10%)	NA	NA	NA	NA	NA
Local	57/215 (27%)	44/103 (43%)	28/31 (90%)	NA	NA	NA	NA	NA
Residential address								
Age, years								
<20	10/215 (5%)	8/93 (9%)	2/41 (5%)	117/449 (26%)	116/1515 (8%)	253/2313 (11%)	6.4% (2.8–12.2)	1.7% (0.2–5.0)
20–59	145/215 (67%)	59/93 (63%)	30/41 (73%)	260/449 (58%)	1086/1515 (72%)	1580/2313 (68%)	18.5% (14.4–23.2)	2.7% (1.8–3.8)
≥60	60/215 (28%)	26/93 (28%)	9/41 (22%)	67/449 (15%)	300/1515 (20%)	462/2313 (20%)	28.0% (19.1–38.2)	2.9% (1.3–5.5)
Sex								
Female	107/215 (50%)	53/93 (57%)	21/41 (51%)	227/449 (51%)	741/1515 (49%)	1149/2313 (50%)	18.9% (14.5–24.0)	2.8% (1.7–4.2)
Male	108/215 (50%)	40/93 (43%)	20/41 (49%)	218/449 (49%)	768/1515 (51%)	1154/2313 (50%)	15.5% (11.3–20.5)	2.5% (1.6–3.9)
Month†								
January	193/215 (90%)	88/93 (95%)	39/41 (95%)	362/449 (81%)	864/1515 (57%)	1546/2313 (67%)	19.6% (16.0–23.5)	4.3% (3.1–5.9)
February	22/215 (10%)	5/93 (5%)	2/41 (5%)	87/449 (19%)	651/1515 (43%)	767/2313 (33%)	5.4% (1.8–12.2)	0.31% (0.04–1.1)
Household size								
≤6 people	188/215 (87%)	79/93 (85%)	32/41 (78%)	309/449 (69%)	1191/1515 (79%)	1799/2313 (78%)	20.4% (16.5–24.7)	2.6% (1.8–3.7)
>6 people	27/215 (13%)	14/93 (15%)	9/41 (22%)	140/449 (31%)	324/1515 (21%)	514/2313 (22%)	9.1% (5.1–14.8)	2.7% (1.2–5.1)
Origin								
Imported	158/215 (73%)	56/93 (60%)	6/41 (15%)	NA	NA	NA	NA	NA
Local	57/215 (27%)	37/93 (40%)	35/41 (85%)	NA	NA	NA	NA	NA

Data are n/N (%) or secondary attack rate (95% CI). When household was defined on the basis of close relatives, the overall data-based secondary attack rates were 13.2% (10.9–15.7) among household contacts and 2.4% (1.6–3.3) among non-household contacts. When household was defined on the basis of residential address, the overall data-based secondary attack rates were 17.2% (14.1–20.6) among household contacts and 2.6% (1.9–3.6) among non-household contacts. Contact type was determined by an individual's relationship with the primary cases of each close contact group. NA=not applicable. *Calculated as the number of secondary cases divided by the sum of secondary cases and non-cases. †Secondary cases and non-cases in each close contact group were allocated to January or February, 2020, on the basis of the number of days in the infectious period of the primary case that occurred in January compared with that in February.

Table 1: Demographic composition of the study population stratified by case type and contact type

which is likely to reflect the tightening of control measures in Guangzhou.

We excluded 12 close contact groups with primary cases but no recorded contact, one close contact group in which all members (two asymptomatic primary cases and two uninfected individuals) stayed in Guangzhou for only 1 day, and 25 contacts with missing data on age or sex. Thus, 182 close contact groups with a total of 332 cases (317 symptomatic and 15 asymptomatic) and 1937 uninfected contacts were included in our transmission analysis. We estimated the secondary attack rates among household and non-household contacts for

the combinations of mean incubation period of 4–7 days and maximum infectious period of 13, 16, and 22 days (appendix 2 p 17). Assessment of the goodness-of-fit for these settings indicated that all models fit the data satisfactorily, because the model-predicted numbers of infections were consistent with the observed values with only small differences before and after the peak of the epidemic (appendix 2 p 30).

Assuming a mean incubation period of 5 days and maximum infectious period of 13 days, when household contact was defined on the basis of close relatives, assuming there was no case isolation, the estimated

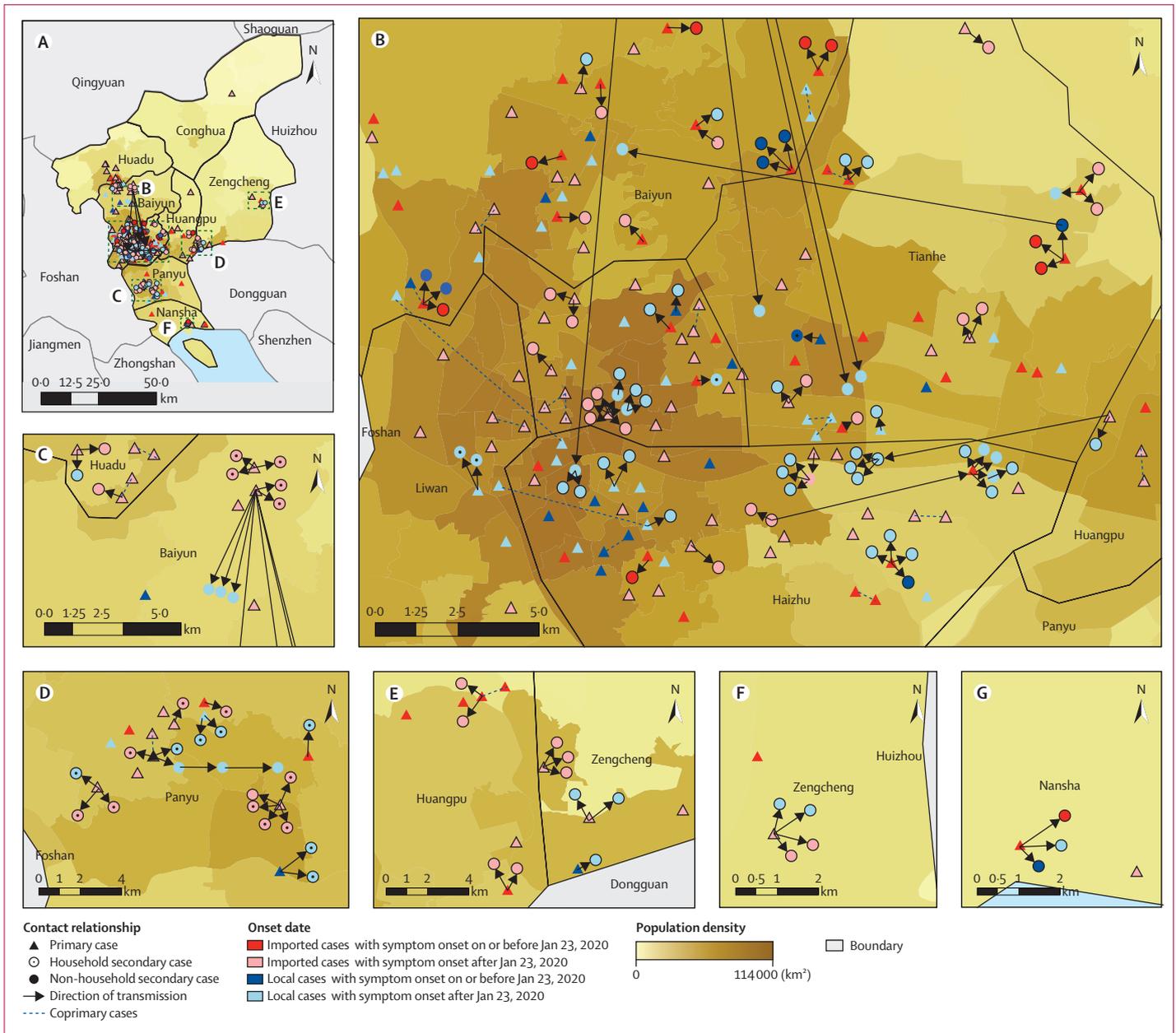


Figure 1: Spatial distribution of COVID-19 case clusters on the basis of contact tracing data from Guangzhou, China, from Jan 7, 2020, to Feb 18, 2020
Overall distribution of COVID-19 case clusters in Guangzhou (A), and distribution in the subregions defined in panel A (B–G). Individuals were considered as primary cases if their symptom onset dates were the earliest or 1 day (≤ 3 days for an imported case) after the earliest in the cluster and as secondary cases otherwise. Non-infected contacts are not shown. The displayed location of each case is randomly perturbed away from the actual residential address.

secondary attack rates were 12.4% (95% CI 9.8–15.4%) among household contacts and 7.9% (95% CI 5.3–11.8%) among non-household contacts. A longer incubation period was associated with a slightly lower secondary attack rate estimate than a shorter incubation period, and a longer infectious period was associated with a slightly higher secondary attack rate estimate than a shorter infectious period (table 2; appendix 2 p 21). When different mean incubation periods and maximum

infectious periods were considered, the secondary attack rate varied from 11.4% (95% CI 9.0–14.2) to 18.0% (13.9–23.0) among household contacts, and from 7.5% (5.0–11.2) to 12.2% (8.0–18.1) among non-household contacts (appendix 2 p 21). The estimated local *R* based on observed contact frequencies of primary cases was 0.5 (95% CI 0.41–0.62), which was insensitive to the assumed incubation and infectious periods. Thus, a typical case infected 0.5 individuals on average in

Guangzhou, implying inefficient transmission of the disease under the control measures. The projected local R_t , had there been no isolation of cases or quarantine of their contacts, was 0.6 (95% CI 0.49–0.74) when household was defined on the basis of close relatives. Higher estimates of projected local R_t were associated with shorter incubation periods and longer infectious periods (appendix 2 p 21). When household was defined on the basis of close relatives, the daily transmission probability during the incubation period was similar to that during the illness period (estimated OR 0.61 [95% CI 0.27–1.38], table 3); however, the difference was much larger when longer incubation periods were assumed (table 3; appendix 2 p 22–23).

Individuals aged 60 years and older were the group most susceptible to SARS-CoV-2 infection (table 4; appendix 2 p 24). Assuming a mean incubation period of 5 days and maximum infectious period of 13 days, in comparison with the oldest age group (≥ 60 years), the risk of infection was lower in the youngest age group (<20 years; OR 0.23 [95% CI 0.11–0.46]) and the 20–59 year age group (0.64 [0.43–0.97]). The person-to-person transmissibility of the virus declined over time to some extent (February vs January OR 0.42 [95% CI 0.17–1.07]). These estimated ORs were insensitive to the assumptions about the natural history of disease (appendix 2 p 24). The estimated probability of daily transmission was two times higher in households of six people or less than in larger households (more than six people; appendix 2 p 25). We found no association between age and infectivity and no associations between sex and susceptibility or infectivity.

Restricting household contacts to those who were living at the same address as the primary case regardless of relationship resulted in higher secondary attack rate estimates among household contacts, ranging from 16.1% (95% CI 12.5–20.4) to 24.3% (18.5–31.2), but lower secondary attack rate estimates among non-household contacts (ranging from 6.8% [5.0–9.2] to 9.3% [6.5–13.1]) under the various settings (ie, mean incubation and maximum infectious periods) of the natural history of disease (appendix 2 p 21). Assuming an incubation time of 5 days and an infectious period of 13 days, the estimated secondary attack rates were 17.1% (95% CI 13.3–21.8) among household contacts and 7.3% (5.4–9.9) among non-household contacts (table 2). The effect of age and the relative infectivity of virus in the illness period versus the incubation period remained similar (table 4; appendix 2 p 24).

Discussion

We retrospectively characterised the spatiotemporal epidemiology and transmissibility of SARS-CoV-2 in Guangzhou, the most populated city in southern China, from early January up to mid-February, 2020. The rapid decline in the R_t indicates the effectiveness of the control policy implemented in the city. Social distancing or other

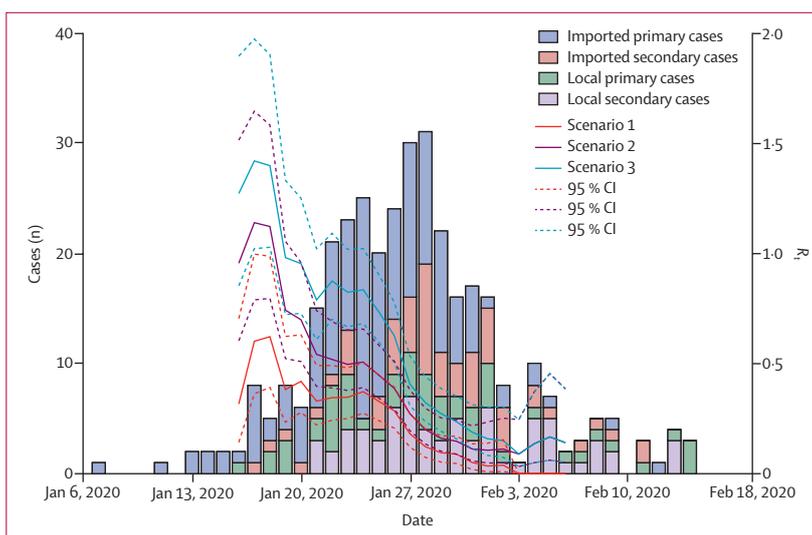


Figure 2: Epidemic curve based on symptom onset dates of COVID-19 cases in Guangzhou from Jan 6, 2020, to Feb 18, 2020

Estimated R_t for three scenarios: scenario 1, all imported cases (who travelled to or resided in Hubei province 14 days before symptom onset) considered as primary cases, and all secondary cases were infected by primary cases in the same case cluster; scenario 2, which is identical to scenario 1, with the additional assumption that local primary cases might have been infected by earlier cases in other clusters; and scenario 3, which is identical to scenario 2, with the additional assumption that imported secondary cases were considered as infected by primary cases in the same cluster. R_t =effective reproductive number.

	Mean incubation period of 5 days		Mean incubation period of 7 days	
	13-day infectious period	22-day infectious period	13-day infectious period	22-day infectious period
Close relatives				
Secondary attack rate, % (95% CI)				
Household	12.4% (9.8–15.4)	15.5% (11.7–20.2)	11.4% (9.0–14.2)	13.1% (9.9–17.1)
Non-household	7.9% (5.3–11.8)	10.4% (6.7–15.8)	7.5% (5.0–11.2)	8.9% (5.7–13.6)
Local R_t (95% CI)				
With quarantine	0.50 (0.41–0.62)	0.51 (0.39–0.66)	0.51 (0.41–0.63)	0.51 (0.39–0.67)
No quarantine	0.60 (0.49–0.74)	0.76 (0.59–1.00)	0.56 (0.45–0.69)	0.65 (0.49–0.85)
Residential address				
Secondary attack rate, % (95% CI)				
Household	17.1% (13.3–21.8)	21.2% (15.8–27.8)	16.1% (12.5–20.4)	18.3% (13.6–24.1)
Non-household	7.3% (5.4–9.9)	9.3% (6.5–13.1)	6.8% (5.0–9.2)	7.8% (5.5–11.0)
Local R_t (95% CI)				
With quarantine	0.50 (0.40–0.61)	0.50 (0.38–0.65)	0.50 (0.41–0.62)	0.51 (0.39–0.66)
No quarantine	0.59 (0.48–0.72)	0.74 (0.57–0.96)	0.55 (0.45–0.67)	0.63 (0.48–0.82)

Estimates were reported using two different definitions of household contact (close relatives or individuals sharing the same residential address) and for selected settings of the natural history of disease. This model was not adjusted for age group, epidemic phase, or household size. R_t =reproductive number.

Table 2: Model-based estimates of secondary attack rates among household and non-household contacts, and local R_t with and without quarantine

potential personal behavioural changes might have also shifted the contact pattern between household members, as shown by the two-fold reduction in the probability of household transmission observed between January and February. Additionally, we assessed the effects of host features and disease stage on susceptibility and

	Mean incubation period of 5 days		Mean incubation period of 7 days	
	13-day infectious period	22-day infectious period	13-day infectious period	22-day infectious period
Close relatives				
Transmission probabilities for household contacts ($\times 10^{-3}$)				
Incubation	1.84 (1.36–2.49)	1.91 (1.44–2.54)	2.09 (1.63–2.69)	2.11 (1.66–2.68)
Illness	1.13 (0.61–2.08)	0.80 (0.44–1.46)	0.54 (0.19–1.57)	0.41 (0.16–1.05)
Transmission probabilities for non-household contacts ($\times 10^{-3}$)				
Incubation	1.16 (0.73–1.83)	1.25 (0.81–1.92)	1.37 (0.9–2.07)	1.4 (0.93–2.1)
Illness	0.71 (0.35–1.43)	0.52 (0.26–1.05)	0.35 (0.11–1.09)	0.27 (0.1–0.75)
Transmission probability from an external source ($\times 10^{-4}$)	1.71 (0.78–3.78)	1.49 (0.65–3.44)	1.54 (0.61–3.86)	1.38 (0.54–3.56)
OR	0.61 (0.27–1.38)	0.41 (0.19–0.89)	0.26 (0.08–0.86)	0.19 (0.07–0.55)
Residential address				
Transmission probabilities for household contacts ($\times 10^{-3}$)				
Incubation	2.64 (1.9–3.66)	2.77 (2.03–3.76)	3.03 (2.29–4.00)	3.07 (2.35–4.02)
Illness	1.58 (0.84–2.95)	1.1 (0.59–2.05)	0.79 (0.28–2.21)	0.57 (0.22–1.46)
Transmission probabilities for non-household contacts ($\times 10^{-3}$)				
Incubation	1.08 (0.75–1.55)	1.14 (0.81–1.61)	1.23 (0.89–1.69)	1.25 (0.92–1.70)
Illness	0.64 (0.33–1.24)	0.45 (0.23–0.87)	0.32 (0.11–0.91)	0.23 (0.09–0.61)
Transmission probability from an external source ($\times 10^{-4}$)	1.74 (0.79–3.84)	1.54 (0.67–3.53)	1.53 (0.6–3.87)	1.4 (0.54–3.62)
OR	0.59 (0.26–1.35)	0.39 (0.18–0.86)	0.26 (0.08–0.82)	0.18 (0.06–0.52)

Data are estimates (95% CI). Estimates of the daily probability of infection from an external source and the ORs for the relative infectivity during the illness versus incubation period are also provided. Estimates are reported using two different definitions of household contact (close relatives, or only individuals sharing the same residential address) and for selected settings of the natural history of disease (ie, mean incubation and maximum infectious periods). This model was not adjusted for age group, epidemic phase, or household size. OR=odds ratio.

Table 3: Model-based estimates of daily transmission probabilities for household contacts and non-household contacts during the incubation and illness periods

infectivity. We found that patients with COVID-19 were at least as infectious in the incubation periods as during their illness periods, and that older people (aged >60 years) are most susceptible to household infection of SARS-CoV-2.

When household contact was defined by residential address, our model-based secondary attack rate estimate for the Guangzhou contact tracing data was 17.1%, which is higher than data-based secondary attack rate estimates of 14.9% for Shenzhen and 10.2% for Guangzhou under the same household definition.^{14,15} Generally, a data-based secondary attack rate estimate would be higher than a model-based secondary attack rate estimate since data-based estimates do not exclude tertiary transmission and untraced exposure. However, these data-based secondary attack rate estimates reflect transmissibility under control measures such as case isolation, whereas our model-based secondary attack rate estimates assumed exposure of a susceptible individual during the whole infectious period of an

	Mean incubation period of 5 days		Mean incubation period of 7 days	
	13-day infectious period	22-day infectious period	13-day infectious period	22-day infectious period
Close relatives				
Susceptibility				
Age <20 years vs ≥ 60 years	0.23 (0.11–0.46)	0.22 (0.11–0.46)	0.22 (0.11–0.45)	0.22 (0.11–0.45)
Age 20–59 years vs ≥ 60 years	0.64 (0.43–0.97)	0.64 (0.42–0.96)	0.63 (0.42–0.95)	0.63 (0.42–0.94)
February vs January	0.42 (0.17–1.07)	0.46 (0.19–1.10)	0.36 (0.12–1.05)	0.38 (0.13–1.09)
Infectivity				
Illness vs incubation period	0.60 (0.27–1.36)	0.42 (0.19–0.91)	0.29 (0.10–0.88)	0.21 (0.07–0.58)
Residential address				
Susceptibility				
Age <20 years vs ≥ 60 years	0.22 (0.11–0.46)	0.22 (0.11–0.45)	0.22 (0.11–0.44)	0.22 (0.11–0.44)
Age 20–59 years vs ≥ 60 years	0.67 (0.45–1.00)	0.67 (0.45–1.00)	0.66 (0.44–0.99)	0.66 (0.44–0.99)
February vs January	0.57 (0.23–1.39)	0.62 (0.27–1.44)	0.50 (0.18–1.36)	0.53 (0.20–1.42)
Infectivity				
Illness vs incubation period	0.54 (0.23–1.26)	0.38 (0.17–0.84)	0.24 (0.07–0.79)	0.18 (0.06–0.52)

Data are OR (95% CI). Estimates were reported for selected settings of the natural history of disease (ie, mean incubation and maximum infectious periods). This model was adjusted for age group, epidemic phase, and household size. OR=odds ratio.

Table 4: Model-based effects of age group and epidemic phase on susceptibility and relative infectivity during the illness period compared with the incubation period

infectior, which is more epidemiologically relevant and generalisable.

Model-based household secondary attack rate estimates for SARS-CoV or the Middle East respiratory syndrome coronavirus (MERS-CoV) are not available; however, a small number of studies have reported data-based secondary attack rate estimates in the household or comparable settings. For SARS-CoV, the secondary attack rate was estimated to be 4.6–8.0% in Beijing, Hong Kong, and Singapore.^{21,22} The daily transmission probabilities during the illness period, however, are comparable between SARS-CoV (0.013 [95% CI 0.011–0.016]) and SARS-CoV-2 (0.016 [0.008–0.029]; appendix 2 p 22).²³ Information about household transmissibility of MERS-CoV is less clear. In a multicity household study in Saudi Arabia, the household secondary attack rate was 4% (95% CI 2–7).²⁴ In an outbreak among 828 female workers who lived in an

expatriate dormitory consisting of 24 villas in Riyadh, Saudi Arabia, 19 workers in seven villas were infected. If each affected villa was considered a household with a mean of 34.5 residents including a single primary case, we estimated a secondary attack rate of 5.1% (95% CI 2.8–9.0).²⁵ We conclude that SARS-CoV-2 is more transmissible than SARS-CoV and MERS-CoV in households.

We quantified the infectivity of patients with COVID-19 during their incubation period using household data, but epidemiological evidence has been published previously. Transmission to secondary cases during the incubation period of the primary case has been reported in Germany and China.^{7,26} An analysis of 77 transmission pairs within and outside of China estimated that nearly half of all transmission events could have occurred during the incubation period.⁵ By contrast, shedding of SARS-CoV peaks 6–11 days after illness onset, and asymptomatic and mild MERS-CoV cases were hypothesised to transmit inefficiently, indicating the importance of symptoms in the transmission of these two coronaviruses.^{27–29} This finding indicates the importance of testing close contacts of COVID-19 cases to identify and isolate infections in the incubation period. Alternatively, the lower estimated transmission probability of SARS-CoV-2 during the illness period than during the incubation period could be partially attributed to self-distancing within households when the primary cases developed symptoms. The infectivity measured by the transmission model accounts for both the biological process of viral shedding and the social contact process, and our data cannot separate the two processes.

We estimated that the local R of SARS-CoV-2 was relatively low (around 0.5), which is consistent with the mean R . We estimated that without isolation of cases or quarantine of their contacts and assuming a mean incubation period of 5 days, the local reproductive number would have been about 20–50% higher, increasing to 0.6–0.76 (appendix 2 p 21). The subcritical local reproductive number ($R < 1$) in the absence of isolation is due to the small average number of contacts per person per day (appendix 2, p 18), which is likely to be a result of the stringent control measures that were implemented in the city. Although the effect of case isolation seems moderate, considering the high infectivity of the virus during the incubation period, quarantine of asymptomatic contacts could have prevented more onward transmissions.

Our analysis has several limitations. The model suggests that a longer assumed incubation period was associated with higher estimated infectivity in this period, which could be due to the rapid quarantine of close contacts after symptom onset of cases. When few transmissions occurred during the illness period, bias could occur due to the paucity of data. Additionally, we were unable to reliably quantify the infectivity of asymptomatic infections, since only two of the

15 asymptomatic infections included in the household transmission analyses were considered primary cases. Some asymptomatic infections might have been missed since close contacts were tested only twice and the tests were done 14 days apart. Moreover, our assumption that asymptomatic infections have the same infectivity as symptomatic cases during their incubation period might not be realistic. Furthermore, it is likely that some imported primary cases might have been infected locally and that some asymptomatic infections or cases might have been missed by contact tracing or by false negative tests, which could lead to underestimation of the R , and the secondary attack rate.

The infectiousness of patients with COVID-19 during their incubation periods is high and could substantially increase the difficulty of curbing the ongoing pandemic. Active case finding and isolation in conjunction with comprehensive contact tracing and quarantine are useful for preventing infected contacts from spreading the virus during their incubation periods, which could be crucial when societal restrictions on human movement and mixing are lifted. The provision of comfortable facilities for exposed contacts to quarantine or for mild cases to isolate away from their families could be a valuable strategy to limit onward transmission within households.

Contributors

QLJ, JY, ZBZ, LL, MMM, YL, YM, and ZCY collected the data. YY, ZCY, and LQF conceived the statistical analysis plan. QLJ, MJL, ARZ, and NJ cleaned the data. MJL, QLJ, ARZ, and NJ did statistical analyses under the supervision of YY, ZCY, and LQF. YY drafted the manuscript. ND, IL, and EK contributed to interpretation of results and writing of the manuscript.

Declaration of interests

We declare no competing interests.

Data sharing

Data sharing requests should be directed to Guangzhou Center for Disease Control and Prevention.

Acknowledgments

This study was supported by grants from the US National Institutes of Health (R01 AI139761, R01 AI116770 and R37 AI32042), the Science and Technology Plan Project of Guangzhou (201804010121), the Project for Key Medicine Discipline Construction of Guangzhou Municipality (2017–2019–04), and the Key Research and Development Program of China (2019YFC1200604). We thank M Elizabeth Halloran (Fred Hutchinson Cancer Research Center, Seattle, WA, USA) for helpful discussions. We also thank the staff members of all district-level Center for Disease Control and Prevention and community health service centres in Guangzhou for their assistance in field investigation and data collection.

References

- 1 Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020; **20**: 533–34.
- 2 Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* 2020; **323**: 1239.
- 3 van Doremalen N, Bushmaker T, Morris DH, et al. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *N Engl J Med* 2020; **382**: 1564–67.
- 4 Xu Y, Li X, Zhu B, et al. Characteristics of pediatric SARS-CoV-2 infection and potential evidence for persistent fecal viral shedding. *Nat Med* 2020; **26**: 502–05.

- 5 He X, Lau EHY, Wu P, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med* 2020; **26**: 672–75.
- 6 Bai Y, Yao L, Wei T, et al. Presumed asymptomatic carrier transmission of COVID-19. *JAMA* 2020; **323**: 1406.
- 7 Rothe C, Schunk M, Sothmann P, et al. Transmission of 2019-nCoV Infection from an asymptomatic contact in Germany. *N Engl J Med* 2020; **382**: 970–71.
- 8 Liu J, Liao X, Qian S, et al. Community transmission of severe acute respiratory syndrome coronavirus 2, Shenzhen, China, 2020. *Emerg Infect Dis* 2020; **26**: 1320–23.
- 9 She J, Jiang J, Ye L, Hu L, Bai C, Song Y. 2019 novel coronavirus of pneumonia in Wuhan, China: emerging attack and management strategies. *Clin Transl Med* 2020; **9**: 19.
- 10 Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet* 2020; **395**: 689–97.
- 11 Chinazzi M, Davis JT, Ajelli M, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* 2020; **368**: 395–400.
- 12 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**: 638–42.
- 13 Liu Y, Eggo RM, Kucharski AJ. Secondary attack rate and superspreading events for SARS-CoV-2. *Lancet* 2020; **395**: e47.
- 14 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; published online April 27. DOI:10.1016/S1473-3099(20)30287-5.
- 15 Luo L, Liu D, Liao X-L, et al. Modes of contact and risk of transmission in COVID-19 among close contacts. *medRxiv* 2020; published online March 26. DOI:10.1101/2020.03.24.20042606 (preprint).
- 16 Burke RM, Midgley CM, Dratch A, et al. Active monitoring of persons exposed to patients with confirmed COVID-19—United States, January–February 2020. *MMWR Morb Mortal Wkly Rep* 2020; **69**: 245–46.
- 17 Yang Y, Longini IM Jr, Halloran ME, Obenchain V. A hybrid EM and Monte Carlo EM algorithm and its application to analysis of transmission of infectious diseases. *Biometrics* 2012; **68**: 1238–49.
- 18 Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med* 2020; **382**: 1199–207.
- 19 Lauer SA, Grantz KH, Bi Q, et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Ann Intern Med* 2020; **172**: 577–82.
- 20 Wölfel R, Corman VM, Guggemos W, et al. Virological assessment of hospitalized patients with COVID-2019. *Nature* 2020; **581**: 465–69.
- 21 Goh DL, Lee BW, Chia KS, et al. Secondary household transmission of SARS, Singapore. *Emerg Infect Dis* 2004; **10**: 232–34.
- 22 Lau JTF, Lau M, Kim JH, Tsui HY, Tsang T, Wong TW. Probable secondary infections in households of SARS patients in Hong Kong. *Emerg Infect Dis* 2004; **10**: 235–43.
- 23 Pitzer VE, Leung GM, Lipsitch M. Estimating variability in the transmission of severe acute respiratory syndrome to household contacts in Hong Kong, China. *Am J Epidemiol* 2007; **166**: 355–63.
- 24 Drosten C, Meyer B, Müller MA, et al. Transmission of MERS-coronavirus in household contacts. *N Engl J Med* 2014; **371**: 828–35.
- 25 Van Kerkhove MD, Alaswad S, Assiri A, et al. Transmissibility of MERS-CoV infection in closed setting, Riyadh, Saudi Arabia, 2015. *Emerg Infect Dis* 2019; **25**: 1802–09.
- 26 Cai J, Sun W, Huang J, Gamber M, Wu J, He G. Indirect virus transmission in cluster of COVID-19 cases, Wenzhou, China, 2020. *Emerg Infect Dis* 2020; **26**: 1343–45.
- 27 Cheng PK, Wong DA, Tong LK, et al. Viral shedding patterns of coronavirus in patients with probable severe acute respiratory syndrome. *Lancet* 2004; **363**: 1699–700.
- 28 Memish ZA, Assiri AM, Al-Tawfiq JA. Middle East respiratory syndrome coronavirus (MERS-CoV) viral shedding in the respiratory tract: an observational analysis with infection control implications. *Int J Infect Dis* 2014; **29**: 307–08.
- 29 Al Hosani FI, Kim L, Khudhair A, et al. Serologic follow-up of Middle East Respiratory Syndrome Coronavirus cases and contacts—Abu Dhabi, United Arab Emirates. *Clin Infect Dis* 2019; **68**: 409–18.