Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Epidemics 39 (2022) 100553

Contents lists available at ScienceDirect

Epidemics

journal homepage: www.elsevier.com/locate/epidemics

Variability in transmission risk of SARS-CoV-2 in close contact settings: A contact tracing study in Shandong Province, China

Tim K. Tsang [a,1], Li-Qun Fang [b,1], Anran Zhang [b,c,1], Fa-Chun Jiang [d,1], Shi-Man Ruan [e,1], Lan-Zheng Liu [e], Benjamin J. Cowling [a,f,*], Wei Liu [b,**], Yang Yang [g,**]

[a] WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China
[b] State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing, China
[c] Department of Epidemiology, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China
[d] Qingdao Municipal Centre for Disease Control and Prevention & Qingdao Institute of Preventive Medicine, Shandong, China
[e] Jinan Municipal Centre for Disease Control and Prevention, Shandong, China
[f] Laboratory of Data Discovery for Health Limited, Hong Kong Science and Technology Park, Hong Kong Special Administrative Region, China
[g] Department of Biostatistics, School of Public Health and Health Professions & Emerging Pathogens Institute, University of Florida, Gainesville, FL 32610, USA

ARTICLE INFO

Keywords:
COVID-19
Transmission risk
Close contact
Households
Superspreading

ABSTRACT

Background: Understanding the relative transmissibility of SARS-CoV-2 virus across different contact settings and the possibility of superspreading events is important for prioritizing disease control. Such assessment requires proper consideration of individual level exposure history, which is made possible by contact tracing.
Methods: The case-ascertained study in Shandong, China including 97 laboratory-confirmed index cases and 3158 close contacts. All close contacts were quarantined after their last exposure of index cases. Contacts were tested for COVID-19 regularly by PCR to identify both symptomatic and asymptomatic infections. We developed a Bayesian transmission model to the contact tracing data to account for different duration of exposure among individuals to transmission risk in different settings, and the heterogeneity of infectivity of cases.
Results: We estimate secondary attack rates (SAR) to be 39% (95% credible interval (CrI): 20–64%) in households, 30% (95% CrI: 11–67%) in healthcare facilities, 23% (95% CrI: 7–51%) at workplaces, and 4% (95% CrI: 1–17%) during air travel. Models allowing heterogeneity of infectivity of cases provided a better goodness-of-fit. We estimated that 64% (95% CrI: 55–72%) of cases did not generate secondary transmissions, and 20% (95% CrI: 15–26%) cases explained 80% of secondary transmissions.
Conclusions: Household, healthcare facilities and workplaces are efficient setting for transmission. Timely identification of potential superspreaders in most transmissible settings remains crucial for containing the pandemic.

1. Introduction

Coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection has claimed more than three million human lives as of May 2021 (Dong et al., 2020). The infection fatality rate was estimated to increase with age, $< 1\%$ for patients aged under 65 and nearly 8% for patients aged 80 + , substantially higher than seasonal influenza (O'Driscoll et al., 2020). Prevention and control of this virus are challenging due to its ability to transmit before symptom onset or from persons without symptoms (Furukawa et al., 2020; Li et al., 2020; Buitrago-Garcia et al., 2020; Tong et al., 2020; Bai et al., 2020). Mass testing, quarantine of cases and isolation of close contacts have been shown effective for containing the spread of the virus, but the required resources are tremendous and not always available (Tong et al., 2020). Multiple vaccines have been showed efficacious in randomized clinical trials and authorized for

* Corresponding author at: WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China.
** Corresponding authors.
E-mail addresses: bcowling@hku.hk (B.J. Cowling), liuwei@bmi.ac.cn (W. Liu), yangyang@ufl.edu (Y. Yang).
[1] These authors contributed equally.

https://doi.org/10.1016/j.epidem.2022.100553
Received 27 September 2021; Received in revised form 3 March 2022; Accepted 7 March 2022
Available online 9 March 2022

emergency use, yet the progress of vaccination has been slow due to both supply shortage and vaccine hesitancy. Several variants of concerns have shown higher transmissibility and possible immune escape, and it may take months to develop and deliver updated vaccines (Bai et al., 2020). It is therefore important to determine the relative transmission efficiency of the virus in different contact settings to prioritize available intervention resources to reach optimal control (Liu et al., 2020).

Transmissibility of SARS-CoV-2 has been studied in several settings including households (Luo et al., 2020; Lewis et al., 2020; Jing et al., 2020; Bi et al., 2020; Madewell et al., 2020; Li et al., 2020c, 2021), healthcare (Luo et al., 2020; Adams and Walls, 2020), and air travel (Bi et al., 2020; Freedman and Wilder-Smith, 2020; Bae et al., 2020). However, most of these studies were either focusing on a single setting, did not take into account crucial confounders such as exposure history of close contacts, or ignored overdispersion in transmissibility across contact groups. Some studies examined the transmissibility specific to multiple contact settings and addressed superspreading events, but non-household settings investigated are often broad rather than specific, e.g., social contacts (Tsang et al., 2020; Lloyd-Smith et al., 2005). Here, we apply an individual-based Bayesian transmission model to the detailed contact tracing data from the first COVID-19 pandemic wave in Shandong Province, China. We aim to estimate the secondary attack rates in different contact settings and to evaluate potential risk factors for infection and transmission.

2. Methods

2.1. Study design

Demographic, clinical and laboratory test data on laboratory-confirmed symptomatic and asymptomatic SARS-CoV-2 infections (index cases) and their close contacts were collected by municipal centers for disease control and prevention (CDC) in Jinan, Jining and Qingdao in Shandong Province in response to the COVID-19 pandemic. Data of clusters identified during 22 January to 30 May, 2020 were retrospectively retrieved from the surveillance database for this study. During this period, newly detected cases were isolated at hospitals and their close contacts were traced. According to the 5th and 6th editions of the COVID-19 Prevention and Control Plan issued by the National Health Commission of China, A close contact of a case is defined as any person who was in close proximity to the case without any personal protection equipment, starting 2 days before the symptom onset of the case or specimen collection if the case was asymptomatic. Close contact settings include but are not limited to (1) living, working, dining or taking classes with the case in the same closed space or in proximity; (2) providing health care to or visiting the case at a hospital; and (3) sharing transportation with and in close proximity to the case (in flights, passengers within 3 rows of seats in the front and back of a case as well as crew members who had been in proximity to a case were considered as close contacts); and (4) other individuals in close proximity to the cases as determined by field investigators.

Each identified close contact was quarantined either at a hospital or a hotel for 14 days. During quarantine, swab specimens were collected at day 1, 4, 7 and 14 for RT-PCR testing. We extracted age, sex, occupation, start and end dates as well as type of contact with others, and start date of quarantine for each participant. The start and end dates of exposure periods were determined from the reported contact history and travel history based on epidemiological investigations, supplemented by data from mobile phone apps, registration records and surveillance cameras. For infected individuals, we additionally extracted information on symptom onset date (for symptomatic infection) or collection date of first test-positive specimens (for asymptomatic infection), and severity status (asymptomatic, mild, moderate, severe, and critical). Severity status was determined by the patient's attending physician based on the Guidelines for Diagnosis and Management of COVID-19 that was issued by the National Health Commission of China (Tsang et al., 2020). All methods used are in compliance with REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) guideline. All experimental protocols for case confirmation were in accordance with the 6th edition of Prevention and Control Plan for COVID-19 issued by the National Health Commission of China (NHCC), which was approved by the Institutional Review Board of NHCC. Written informed consent was waived by the National Health Commission of China for health data collected during outbreak investigations of notifiable infectious diseases.

2.2. Case definitions

Following the definition in the Guidelines for Diagnosis and Treatment of COVID-19 published by the National Heath Commission, a symptomatic COVID-19 case refers to a laboratory-confirmed SARS-CoV-2 infection presenting at least two of the clinical signs (Dong et al., 2020): Fever and/or respiratory symptoms (O'Driscoll et al., 2020); Radiographic characteristics of pneumonia, such as multiple ground-glass shadows, infiltrative shadows and consolidation in both lungs (Furukawa et al., 2020); normal or lower leukocyte counts, or lower lymphocyte counts at acute phased of the disease. During the first wave, laboratory-confirmation was done uniformly with RT-PCR on nasal swaps. An asymptomatic SARS-CoV-2 infection was defined as a test-positive individual without clinical manifestation throughout the course of infection. Asymptomatic infections were identified mostly via contact-tracing. A close contact was defined as an individual who had unprotected contact within 1 m with a suspected or confirmed case within 2 days before symptom onset or, if the case was asymptomatic, collection of the first test-positive specimen.

While contact tracing was initiated by discovery of an index case, the traced close contact group may contain multiple cases among whom the index case was not necessarily the earliest case. For each cluster, we treat as day 1 the date of the earliest symptom onset (symptomatic infection) or the first test-positive specimen (asymptomatic infection). A primary case was defined as the case who had the earliest symptom onset (symptomatic) or test-positive specimen (asymptomatic) in a cluster. A close contact group may have multiple primary cases to whom we also refer as co-primary cases. Cases who were not the primary case were classified as secondary cases.

2.3. Statistical analysis

Demographic characteristics and clinical outcomes (if applicable) were summarized for primary and secondary cases as well as their close contacts separately. We used maximum likelihood method to fit geometric and negative binomial distributions to both the number of close contacts per case (primary or secondary) and the number of secondary cases per primary case, potential overdispersion in the observed frequencies will be captured if the negative binomial with mean μ and overdispersion parameter k fits the data better (Lloyd-Smith et al., 2005). The goodness-of-fit was assessed using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Smaller AIC and BIC indicated a better model fit, and difference > 5 was considered as substantial improvement. An important measure for transmissibility of a pathogen is the secondary attack rate (SAR), defined as the probability that a susceptible person is infected by an infectious person via close contact during the infector's whole infectious period. We first calculated a crude SAR as the average proportion of secondary cases among close contacts across all close contact groups with a single primary case and refer to it as the data-based SAR, assuming all secondary cases were infected by their primary case. As a sensitivity analysis, we also calculated data-based SAR using all clusters including those with multiple primary cases.

Not all secondary cases were actually infected by the primary cases, as there could be tertiary transmissions, or infection acquired from outside cluster. In addition, information on symptom onset dates and

detailed contact history inform us about the individual level exposure history, which will improve the estimation of the SAR and effects of associated risk factors (Yang et al., 2006). To address these issues, we developed an individual-based Bayesian model to describe daily transmission dynamics among cases and their close contacts (Appendix, Section 2). Briefly, the daily probability of infection depends on the force of infection from non-specific sources in the community, and contact settings. Mathematically, the probability of infection at day t for individual j takes a chain-binomial form:

$$p_j(t) = 1 - (1-b) * \prod_i (1 - p_{ij}(t)),$$

where b is the infection probability from non-specific sources in the community, and $p_{ij}(t)$ is the covariate-adjusted pairwise probability of transmission from individual i to j at day t. we let $p_{ij}(t)$ depend on covariates via a logistic regression:

$$\text{logit}\left(p_{ij}(t)\right) = \text{logit}\left(\prod_{k=1}^{5} {p_k}^{\delta_{ij}(k)}\right) + \boldsymbol{\beta}' \boldsymbol{X_{ij}}(t) + \beta_{i0},$$

where $p_k, k = 1, \ldots, 5$, are the baseline daily person-to-person transmission probabilities specific to five contact settings. (1 =households, 2 =healthcare facilities, 3 =workplaces, 4 =air transportation and 5 =other), $\delta_{ij}(k)$ indicates the contact setting between i and j, and the vectors $\boldsymbol{X_{ij}}(t)$ and $\boldsymbol{\beta}$ encode the covariates affecting susceptibility or infectivity and associated coefficients.

The covariates considered in this study include age group (0–19, 20–39, 40–59, 60 +), sex, city (Jinan, Jining, Qingdao), and occupation (medical personnel or not) of each close contact, as well as the severity level of each case (asymptomatic or mild, moderate, severe or critical) and the time-dependent symptom status (during incubation period vs. during illness). The last term, $\beta_{i0} \sim N\left(0, \frac{1}{\tau}\right)$ if i is a primary case and $\beta_{i0} = 0$ otherwise, is a random effect added to account for heterogeneity (both overdispersion and zero-inflation) in the infectivity of the primary cases. This random effect model was compared to the model without random effects and assuming homogeneous baseline infectivity (before covariate adjustment) among primary cases. The distributions of the incubation period and the illness period were derived from previous publications (Jing et al., 2020; Li et al., 2020a, 2021; Lauer et al., 2020). We considered different combinations (Table S1) of two possible mean durations of the incubation period (5 and 7 days) and two maximum durations of the infectious period (13 and 21 days), and used results based on a mean incubation period of 5 days and a maximum infectious period of 21 days as the primary results (Jing et al., 2020). We allowed the infectiousness of each case to differ between the incubation period and the illness period, a feature known for SARS-CoV-2 (Li et al., 2021). For each asymptomatic case, we assumed the infectious period started from the day of infection with constant infectivity and was distributed according to the convolution of the incubation and illness periods of symptomatic cases.

All models were fitted using a data-augmented Markov chain Monte Carlo algorithm. All parameters and missing data on age and sex among cases and close contacts were sampled within the Bayesian framework (Appendix Section 3). We evaluated the model's goodness-of-fit by simulating epidemics among the close contact groups with given primary cases using parameters sampled from the model-estimated posterior distributions, and then comparing the simulated temporal distribution of illness onsets of secondary cases with the observed one. We also compared the model-expected number of secondary cases with the observed one for each cluster (Appendix Section 3.5). All statistical analyses were conducted using the statistical platform R version 4.0.5 (R Foundation for Statistical Computing, Vienna, Austria). Code is available at Github: https:// github.com/timktsang/COVID_shandong.

3. Results

Between January 22, 2020 and May 30, 2020, a total of 199 cases were reported and their close contacts were traced by the municipal centers for disease control and prevention of Jinan, Jining and Qingdao, forming 89 unrelated close contact groups with 97 primary cases (8 clusters with 2 co-primary cases). Eight clusters with a single primary case had no contacts, and there were 3158 close contacts in the remaining 81 clusters. The median size of these 89 close contact groups was 23 (IQR: 6–50, range 1–224). We fitted both geometric and negative binomial distributions to the observed numbers of close contacts per case, stratified by contact setting. The negative binomial distribution provided a better fit for all contact settings, with the estimated overdispersion parameter ranging from 0.01 to 0.34, confirming the presence of overdispersion in the number of close contacts (Fig. 1; Table S2). Assuming that all secondary cases were infected by primary cases, on average, a primary case generated 1.05 secondary cases overall in Shandong Province, and the observed reproductive number varied by contact setting: 0.70, 0.14, 0.08 and 0.2 for households, healthcare facilities, workplaces and aircrafts, respectively. Combining all settings together, the negative binomial distribution also fits the numbers of secondary cases per primary case better (Fig. 2; Table S2), with the overdispersion parameter estimated as 0.25. Notably, about 66% (64/97) of the primary cases did not generate any secondary case and the corresponding fitted probability is 0.72, indicating the presence of zero-inflation, that is, the absence of any secondary case in some close contact groups occurred more frequently than a Poisson distribution (k is infinity) or a geometric distribution (k = 1) can explain. In addition, the AIC and BIC values for the negative binomial model was much smaller than those for the geometric model with differences > 20, confirming the superiority of the former in fitting the data.

The mean (IQR) age of primary cases was 37 [30–53] years, similar to that of their close contacts (Table 1). Secondary cases were slightly older than primary cases, mean(IQR)= 41(28−57) years. There were more female than male (65% vs. 35%) primary cases but more male than female (59% vs. 41%) secondary cases. Primary cases were more likely to be severe or critical (14% vs. 5%, p = 0.02) and less likely to be asymptomatic (6% vs. 10%, p = 0.04) than secondary cases. The overall data-based SAR was 3.53% (95% confidence interval (CI): 2.86–4.30%). The data-based SAR was the highest among close contacts aged 60 + years, 8.46% (5.44–12.42%), followed by 6.25% (3.75–9.70%) in children and teenagers < 20 years old, and the lowest data-based SAR was seen in young adults 20–39 years old. Among all the contact settings, household was associated with the highest data-based SAR, 10.1% (95% CI: 7.9–12.6%), and air transportation was associated with the lowest, 0.43% (95% CI: 0.05–1.54%). The data-based SAR computed without excluding the clusters with multiple primary cases were similar (Table S3).

Applying a Bayesian transmission model to the contact-tracing data with individual-level exposure details, we estimated the infection risk per daily exposure and the effects of potential risk modifiers. In total 81 close contacts groups with 89 primary/co-primary cases and 3158 close contacts were included in the transmission modeling analysis (Fig. S1; Table S4). We first estimated the daily transmission probabilities under each contact setting during the incubation period and during illness (Table 2). We estimated that the daily transmission probability of infected individuals to their susceptible close contacts during their incubation period was 0.044 (95% credible interval [CrI]: 0.020, 0.100) within households, 0.032 (95% CrI: 0.009, 0.011) in healthcare facilities, 0.023 (95% CrI: 0.007, 0.071) at workplaces, 0.004 (95% CrI: 0.001–0.016) during air travel, and 0.002 (95% CrI: 0.001, 0.005) in all other settings. The corresponding model-based SAR estimates were 0.39 (95% CrI: 0.20, 0.64), 0.3 (95% CrI: 0.11, 0.67), 0.23 (95% CrI: 0.07, 0.51), 0.04 (95% CrI: 0.01, 0.17) and 0.02 (95% CrI: 0.01, 0.05) for households, healthcare facilities, workplaces, air transportation and other settings, respectively (Table S5).

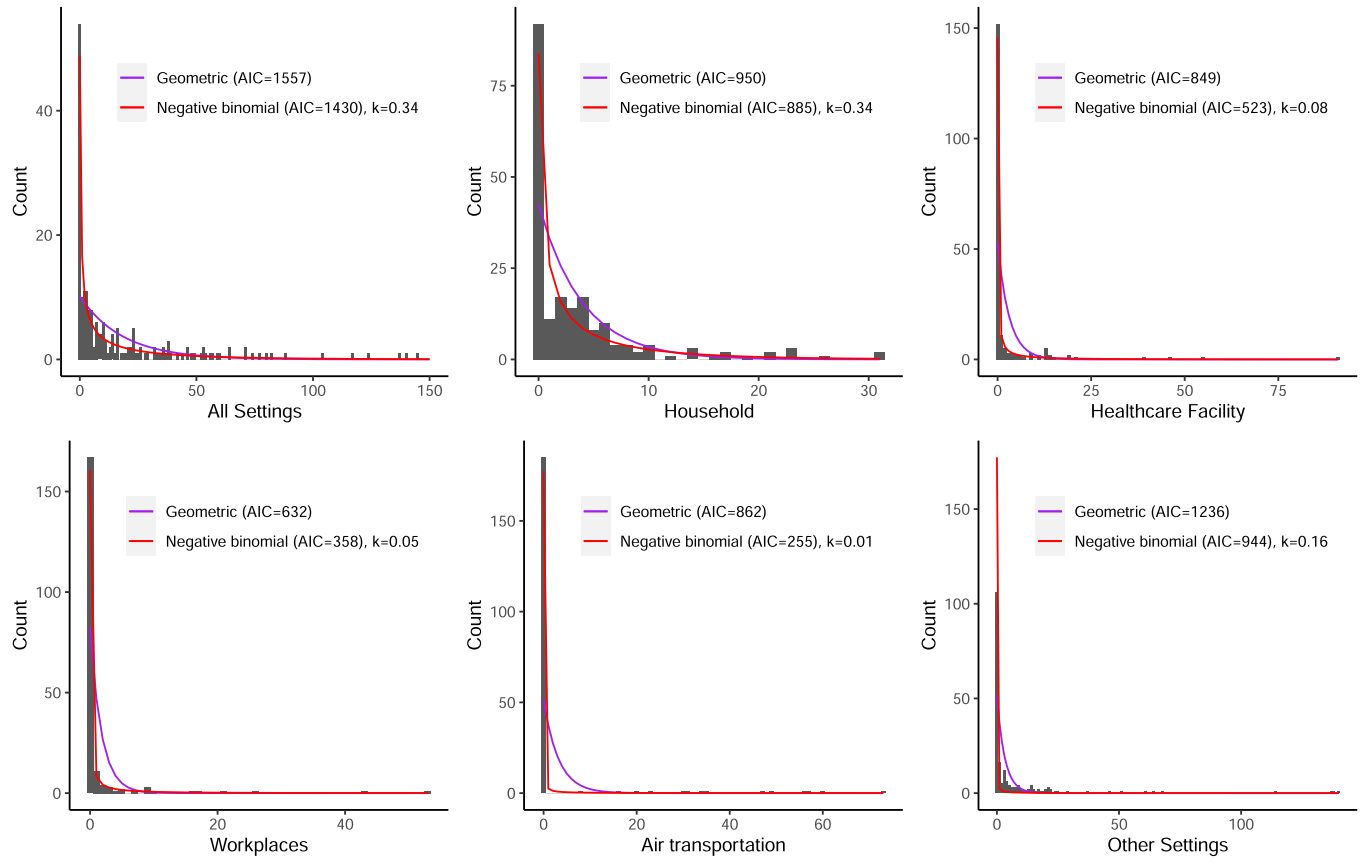

Fig. 1. The distribution for number of close contacts per each case overall and by contact type. The curve indicated the expected number from fitted distribution.

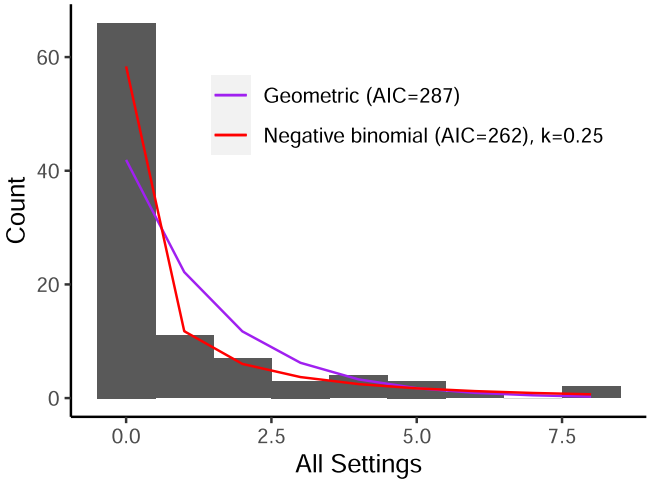

Fig. 2. The distribution for number of secondary cases per each primary case. The curve indicated the expected number from fitted distribution.

Daily transmission probabilities were slightly lower during the illness period, but the difference was moderately sensitive to the assumption about the natural history of disease. When either a longer incubation period or a longer infectious period was assumed, the estimated transmission probabilities during the incubation period increased moderately, whereas those during the illness period decreased, and the estimated relative infectivity of the incubation period vs. the illness period also increased (Table 2). This estimated relative infectivity was the highest when both the incubation period and the infectious period were assumed long. The model-based SAR estimates ranged 23%−39% in households, 17%−30% in healthcare facilities, 14%−23% at workplaces, and 1%−4% during air transportation or in other settings (Table S5).

In this multivariable analysis (Table 3), close contacts younger than 60 had 36–49% lower odds of infection than those aged 60 + (Table 3). Medical personnel were 65% (95% CrI: 10–89%) less likely to be infected than non-medical contacts. Compared to within households, the odds of infection was much lower during air travel, OR= 0.08 (95% CrI: 0.01, 0.34), and in other settings, OR= 0.04 (95% CrI: 0.01, 0.09). The risk of infection in healthcare facilities and workplaces were slightly lower, with odds ratios of 0.73 and 0.52 respectively. Close contacts in Jining and Qingdao were at 79% (95% CrI: 49%, 90%) and 59% (95% CrI: −2%, 81%) lower odds of infection compared with those in Jinan. No impact was found for the severity level of primary cases. Most estimates were robust to the assumption about the incubation and infectious periods.

Using model-estimated parameters, we simulated transmission dynamics among close contacts of primary cases (Fig. S2). The models under different assumptions about the incubation and infectious periods share a similar shape in the average trend over all simulated epidemics that aligned well with the observed epidemic, suggesting (Dong et al., 2020) the models fit the data satisfactorily, and (O'Driscoll et al., 2020) the data do not contain sufficient information about the natural history of disease. In the simulated data based on the model assuming a mean incubation of 5 days and a maximum infectious period of 22 days, 64% (95% CrI: 55%, 72%) of cases did not generate secondary transmissions, and 20% (95% CrI: 15%, 26%) cases explained 80% of secondary transmissions.

The model-predicted total numbers of secondary cases were close to the observed ones for most close contact groups (Fig. 3), further assuring the goodness-of-fit. Two households and one healthcare facility, each with a single primary case, were separated from other close contact

Table 1
Demographic information of primary/co-primary and secondary cases.

	Primary or co-primary cases	Secondary cases	All close contact	Data-based secondary attack rate (%)§
Overall	97	102	3158	3.53 (2.86–4.30)
Age, mean (IQR)	37 [30,53]	41 (28, 57)	36 (26, 50)	
Age groups				
0–19	7/97 (7%)	18/102 (18%)	315/2738 (12%)	6.25 (3.75–9.70)
20–39	45/97 (46%)	29/102 (28%)	1261/2738 (46%)	2.60 (1.72–3.75)
40–59	35/97 (36%)	32/102 (31%)	854/2738 (31%)	3.96 (2.63–5.72)
60+	10/97 (10%)	23/102 (23%)	308/2738 (11%)	8.46 (5.44–12.42)
Sex				
Female	63/97 (65%)	42/102 (41%)	1543/2950 (52%)	3.09 (2.19–4.24)
Male	34/97 (35%)	60/102 (59%)	1407/2950 (48%)	4.50 (3.44–5.78)
Medical personnel				
Yes	0/25 (0%)	5/74 (7%)	333/2632 (13%)	1.56 (0.51–3.61)
No	25/25 (100%)	69/74 (93%)	2299/2632 (87%)	3.35 (2.58–4.27)
Location				
Jinan	23/97 (24%)	28/102 (27%)	540/3158 (17%)	5.23% (3.51–7.48)
Jining	23/97 (24%)	34/102 (33%)	1237/3158 (39%)	3% (2–4.31)
Qingdao	51/97 (52%)	40/102 (39%)	1381/3158 (44%)	3.18% (2.27–4.33)
Contact type$				
Household	67/97 (69%)	68/102 (67%)	674/3158 (21%)	11.07% (8.67–13.87)
Medical-related facilities	24/97 (25%)	14/102 (14%)	504/3158 (16%)	2.89% (1.59–4.8)
Workplace	24/97 (25%)	8/102 (8%)	244/3158 (8%)	2.5% (0.82–5.74)
Air transportation	14/97 (14%)	2/102 (2%)	539/3158 (17%)	0.43% (0.05–1.54)
Other	60/97 (62%)	10/102 (10%)	1197/3158 (38%)	0.85% (0.37–1.66)
Highest severity among primary or co-primary case*				
Asymptomatic or mild	19/97 (20%)	8/102 (8%)	281/3158 (9%)	3.01% (1.31–5.84)
Moderate	63/97 (65%)	62/102 (61%)	2199/3158 (70%)	3.08% (2.34–3.98)
Severe or critical	15/97 (15%)	32/102 (31%)	678/3158 (21%)	5.08% (3.48–7.14)
Severity of individual infections				
Asymptomatic	6/97 (6%)	10/102 (10%)		
Mild	16/97 (16%)	19/102 (19%)		
Moderate	61/97 (63%)	68/102 (67%)		
Severe	12/97 (12%)	4/102 (4%)		
Critical	2/97 (2%)	1/102 (1%)		

$ Primary/co-primary cases may have multiple types of contacts, so the numbers of primary cases do not sum to 97.
*For cluster with co-primary cases, this variable was defined according to the most severe primary case.
§ Observed SAR was calculated based on 73 close contact groups with a single primary case (no co-primary case).

groups in terms of large (≥7) observed and predicted numbers of secondary cases, implying the possibility superspreading events in the three close contact groups. When the random effects were removed from the model, the differences between model-predicted and observed numbers of secondary cases substantially increased (Fig. S3), affirming the necessity of the random effects to account for the overdispersion in individual-level transmissibility.

4. Discussion

In this study, we assessed the transmissibility of SARS-CoV-2 in different settings of contact based on contact tracing data in Shandong Province of China. Regarding transmission between an infectious person and a susceptible contact, SARS-CoV-2 was more transmissible within households than in workplaces and healthcare facilities, and the transmission risk was small during air travel or in other contact settings. In addition to contact setting, age, medical occupation and residential city were risk modifiers.

The higher transmissibility of SARS-CoV-2 within households compared to other contact settings is consistent with findings in other studies (Luo et al., 2020; Jing et al., 2020; Madewell et al., 2020; Li et al., 2021). A possible reason is that interactions within households often feature longer time, closer distance and lack of protection by facial masks, compared to non-household settings (Adam et al., 2020; Sun et al., 2020). In addition to the highest SAR, household was also associated with the highest number of secondary cases generated per primary case, as primary cases most likely had made close contact in their households but not necessarily in other settings. Meanwhile, a primary case may generate on average comparable secondary cases at workplace or in a healthcare facility, compared with in households. The importance of transmission at workplace or in a healthcare facility should also be considered.

The estimated household SARs in Shandong, 23–39% under different assumptions of the natural history of disease were higher than most household transmission studies in China (Jing et al., 2020; Madewell et al., 2020; Li et al., 2021). These SAR estimates were calculated for the whole infectious period, not limited to the individual duration of exposure and therefore eliminating the impact of case isolation and quarantine of close contacts, which is likely more generalizable (Jing et al., 2020; Li et al., 2021). The data-based household SAR, 11%, can be interpreted as the effective transmissibility of the virus under the implemented case isolation and quarantine of close contacts but does not reflect the full potential without them and is therefore not generalizable. On the other hand, the differences between the model-estimated SARs and the data-based SARs clearly demonstrated the effectiveness of the case isolation and quarantine of close contacts. The relative reductions in model-estimated SARs compared to

Table 2
Model-based estimates of daily transmission probabilities from models with a single covariate for relative infectivity of the incubation period vs. the illness period.

Mean incubation period	5 days		7 days	
Duration of infectious period	13 days	22 days§	13 days	22 days
From community (10–4)	6.85 (2.39, 16.88)	3.34 (0.71, 10.08)	6.61 (2.26, 15.9)	3.3 (0.68, 10.02)
During incubation period (10–2)				
Household	3.64 (1.56, 7.8)	4.37 (1.97, 10)	4.54 (1.8, 10.49)	5.72 (2.48, 12.26)
Healthcare facilities	2.1 (0.52, 7.25)	3.23 (0.94, 10.9)	3.32 (0.85, 11.34)	4.74 (1.34, 15.71)
Workplace	1.53 (0.39, 4.58)	2.3 (0.71, 7.08)	1.96 (0.48, 7.1)	3.03 (0.92, 9.71)
Air transportation	0.32 (0.05, 1.32)	0.36 (0.05, 1.62)	0.32 (0.04, 1.7)	0.4 (0.06, 1.98)
Others	0.12 (0.02, 0.39)	0.17 (0.05, 0.5)	0.17 (0.04, 0.57)	0.23 (0.07, 0.65)
During illness (10–2)				
Household	3.57 (1.75, 7.6)	2.96 (1.37, 6.08)	2.56 (0.97, 5.82)	2.62 (1.01, 5.89)
Healthcare facilities	2.14 (0.61, 5.78)	2.16 (0.73, 6.51)	1.84 (0.48, 5.75)	2.14 (0.66, 6.64)
Workplace	1.49 (0.42, 4.45)	1.57 (0.47, 4.25)	1.08 (0.27, 3.8)	1.38 (0.4, 4.26)
Air transportation	0.31 (0.05, 1.45)	0.24 (0.03, 1.13)	0.17 (0.02, 1.04)	0.19 (0.02, 0.99)
Others	0.12 (0.02, 0.34)	0.11 (0.03, 0.29)	0.09 (0.02, 0.29)	0.1 (0.03, 0.27)
Odds ratio for infectivity, incubation period vs. illness period	1.01 (0.46, 2.21)	1.49 (0.76, 3.07)	1.77 (0.81, 4.21)	2.21 (1.08, 4.66)

§ This column is presented as the primary result.

Table 3
Model-estimated odds ratios and 95% credible intervals for potential risk factors for transmission of SARS-CoV-2 in Shandong Province, China.

Mean incubation period	5 days		7 days	
Duration of infectious period	13 days	21 days§	13 days	21 days
Age of contact (year)				
0–19	0.54 (0.26, 1.10)	0.55 (0.26, 1.16)	0.55 (0.26, 1.11)	0.56 (0.25, 1.21)
20–39	0.49 (0.28, 0.91)	0.51 (0.27, 0.96)	0.49 (0.26, 0.97)	0.50 (0.26, 0.98)
40–59	0.65 (0.36, 1.17)	0.64 (0.35, 1.20)	0.66 (0.36, 1.24)	0.63 (0.34, 1.20)
60+	Ref	Ref	Ref	Ref
Sex of contact				
Male vs. Female	1.16 (0.75, 1.82)	1.16 (0.75, 1.84)	1.22 (0.78, 1.94)	1.19 (0.77, 1.89)
Contact was medical personnel				
Yes vs. no	0.41 (0.13, 1.11)	0.35 (0.11, 0.9)	0.41 (0.11, 1.15)	0.35 (0.11, 0.99)
Contact Setting				
Households	Ref	Ref	Ref	Ref
Healthcare facilities	0.59 (0.21, 1.53)	0.73 (0.29, 1.73)	0.74 (0.23, 1.85)	0.81 (0.26, 1.91)
Workplace	0.40 (0.13, 1.01)	0.52 (0.19, 1.29)	0.43 (0.13, 1.18)	0.55 (0.19, 1.46)
Air transportation	0.09 (0.01, 0.36)	0.08 (0.01, 0.34)	0.08 (0.01, 0.34)	0.08 (0.01, 0.33)
Others	0.03 (0.01, 0.08)	0.04 (0.01, 0.09)	0.04 (0.01, 0.10)	0.04 (0.01, 0.10)
Severity of primary case				
Asymptomatic or mild	0.48 (0.12, 1.48)	0.51 (0.11, 1.71)	0.52 (0.13, 1.64)	0.48 (0.12, 1.69)
Moderate	Ref	Ref	Ref	Ref
Severe or critical	1.18 (0.48, 2.71)	1.06 (0.43, 2.65)	1.03 (0.43, 2.46)	1.0 (0.35, 2.63)
Location				
Jinan	Ref	Ref	Ref	Ref
Jining	0.26 (0.12, 0.54)	0.21 (0.1, 0.51)	0.26 (0.12, 0.54)	0.2 (0.08, 0.47)
Qingdao	0.43 (0.18, 0.99)	0.41 (0.16, 1.02)	0.45 (0.19, 1.02)	0.37 (0.15, 1.03)
Relative infectivity				
Incubation vs. illness	1.0 (0.48, 2.03)	1.47 (0.7, 3.22)	1.76 (0.76, 4.07)	2.32 (1.02, 4.69)

§ This column is presented as the primary result.

data-based SARs were 71.6%, 90.3%, 89.1%, 89.3% and 57.5% for household, healthcare facility, workplace, air transportation and other settings, respectively.

Our results also emphasized on the importance of adjusting transmission analyses for individual-level exposure history, which could be different among individuals due to case isolation and quarantine of close contacts. For example, the data-based SAR within households was about 4-fold of those in healthcare facilities and workplaces (Table 1), whereas the model-estimated SARs suggested less than 2-fold differences (Table S3).

Heterogeneous transmissibility of SARS-CoV-2 across close contact groups is a common phenomenon. The number of contacts per case in Shandong Province was more overdispersed ($k = 0.34$) than observed in Hunan province ($k = 0.72$) (Lloyd-Smith et al., 2005). The number of secondary cases per primary case in Shandong was similarly overdispersed ($k = 0.25$) compared to Hong Kong ($k = 0.33$), but more overdispersed than Shenzhen ($k = 0.58$) (Adam et al., 2020). We showed via simulation that 80% of secondary transmissions were generated by 20% cases, also in accordance with previous findings (Adam et al., 2020; Sun et al., 2020). Our study is unique, however, in that we identified possible superspreading events in both households and healthcare facilities, whereas prior studies mostly found such events in non-household or non-healthcare settings (Adam et al., 2020; Lin et al., 2020; Shim et al., 2020).

Our study has several limitations. First, we relied on symptom onset and confirmation dates as well as self-reported exposure history to determine primary cases, but we cannot rule out possible misclassification of primary cases. Second, we assumed that the infectiousness of asymptomatic cases was the same as the symptomatic cases in their incubation period, which may be incorrect (Madewell et al., 2020; Li et al., 2021). However, the number of asymptomatic infections was small and unlikely to affect major results. Third, while contact tracing in China was technology-aided (e.g., mobile phone tracking and public surveillance cameras) and thus relatively comprehensive, it was possible that contact tracing was biased towards acquaintances especially in the early phase of the first wave. Consequently, secondary attack rates in some settings could have been overestimated. Moreover, certain vulnerable subpopulations (e.g., elderly people) might be more likely to be tested, potentially leading to biased results. Finally, asymptomatic infections might have been under-detected, as the proportion of asymptomatic infections in our study was lower than other studies (Madewell et al., 2020; Sun et al., 2020; Davies et al., 2020).

In conclusion, SARS-CoV-2 was mostly efficiently transmitted within households, and its transmissibility in healthcare facilities and workplaces is lower but appreciable. Non-pharmaceutical interventions and vaccination should target large close contact groups in these settings. While pairwise transmission risk was not as strong during air travel, the large number of exposed flight passengers and the implication for long-distance dissemination warrants prevention and control efforts as well. Given that many people are still vaccine-hesitant and viral mutants may

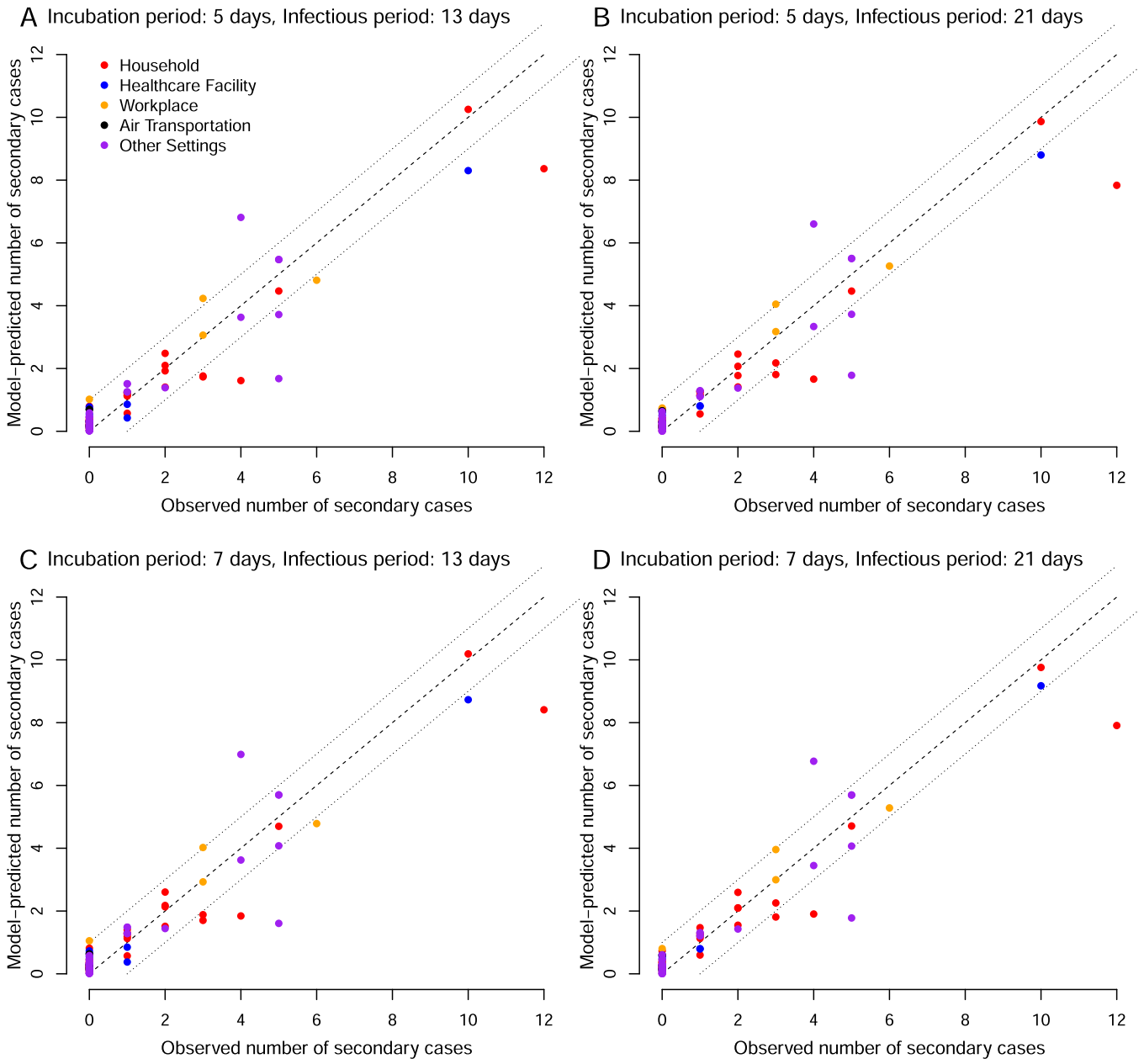

Fig. 3. Scatter plots of model-predicted vs. observed number of secondary cases in each close contact group, by assumption about the mean incubation period and the maximum infectious period. The predicted numbers are from the models include individual-level random effects. Closeness of dots to the diagonal dashed line indicates agreement between model prediction and observation. Dotted lines serve as error band for a difference of ≤ 1 between model-predicted and observed numbers.

escape immunity, timely identification of potential superspreaders and their contacts, whenever feasible, remains crucial for containing the pandemic.

CRediT authorship contribution statement

Tim K. Tsang: Conceptualization, Methodology, Software, Data curation, Writing – original draft, Visualization, Validation. **Li-Qun Fang**: Conceptualization, Methodology, Data curation, Investigation. **Anran Zhang**: Data curation, Investigation. **Fa-Chun Jiang**: Data curation, Investigation. **Shi-Man Ruan**: Data curation, Investigation. **Lan-Zheng Liu**: Data curation, Investigation. **Benjamin J. Cowling**: Conceptualization, Methodology, Supervision, Writing – review & editing. **Wei Liu**: Conceptualization, Methodology, Supervision. **Yang Yang**: Conceptualization, Methodology, Data curation, Writing – original draft, Supervision, Validation, Writing – review & editing.

Acknowledgements

We thank the staff member of all district-level CDCs and community health service centers in Shandong for their assistance in field investigation and data collection. We thank Can Wang for technical assistance. YY was supported by the US National Institutes of Health (grant R56 AI148284) and the US National Science Foundation (grant 2034364). BJC and TKT were supported by the Health and Medical Research Fund, Food and Health Bureau, Government of the Hong Kong Special Administrative Region (grant no. COVID190118) and the Collaborative Research Fund (Project No. C7123-20G) of the Research Grants Council of the Hong Kong SAR Government. BJC is supported by the AIR@-innoHK program of the Innovation and Technology Commission of the

Hong Kong SAR Government. LQF was supported by the National Key Research and Development Program of China (grant 2019YFC1200604) and Beijing Association for Science and Technology (grant 2021-1G-4271).

Conflicts of interest

BJC reports honoraria from Sanofi Pasteur, GSK, AstraZeneca, Moderna and Roche. All other authors report no potential conflicts of interest.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.epidem.2022.100553.

References

Adam, D.C., Wu, P., Wong, J.Y., Lau, E.H.Y., Tsang, T.K., Cauchemez, S., et al., 2020. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. Nat. Med. 26 (11), 1714–1719.
Adams, J.G., Walls, R.M., 2020. Supporting the health care workforce during the COVID-19 global epidemic. JAMA 323 (15), 1439–1440.
Bae, S.H., Shin, H., Koo, H.Y., Lee, S.W., Yang, J.M., Yon, D.K., 2020. Asymptomatic transmission of SARS-CoV-2 on evacuation flight. Emerg. Infect. Dis. 26 (11), 2705–2708.
Bai, Y., Yao, L., Wei, T., Tian, F., Jin, D.Y., Chen, L., et al., 2020. Presumed asymptomatic carrier transmission of COVID-19. JAMA 323 (14), 1406–1407.
Bi, Q., Wu, Y., Mei, S., Ye, C., Zou, X., Zhang, Z., et al., 2020. 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. 20 (8), 911–919.
Buitrago-Garcia, D., Egli-Gany, D., Counotte, M.J., Hossmann, S., Imeri, H., Ipekci, A.M., et al., 2020. Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: a living systematic review and meta-analysis. PLoS Med. 17 (9), e1003346.
Davies, N.G., Klepac, P., Liu, Y., Prem, K., Jit, M., group CC-w, 2020. Age-dependent effects in the transmission and control of COVID-19 epidemics. Nat. Med. 26 (8), 1205–1211.
Dong, E., Du, H., Gardner, L., 2020. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect. Dis. 20 (5), 533–534.
Freedman, D.O., Wilder-Smith, A., 2020. In-flight transmission of SARS-CoV-2: a review of the attack rates and available data on the efficacy of face masks. J. Travel Med. 27, 8.
Furukawa, N.W., Brooks, J.T., Sobel, J., 2020. Evidence supporting transmission of severe acute respiratory syndrome coronavirus 2 while presymptomatic or asymptomatic. Emerg. Infect. Dis. 26 (7), e201595.
Jing, Q.L., Liu, M.J., Zhang, Z.B., Fang, L.Q., Yuan, J., Zhang, A.R., 2020. Household secondary attack rate of COVID-19 and associated determinants in Guangzhou, China: a retrospective cohort study. Lancet Infect. Dis. 20 (10), 1141–1150.
Lauer, S.A., Grantz, K.H., Bi, Q., Jones, F.K., Zheng, Q., Meredith, H.R., 2020. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann. Intern. Med. 172 (9), 577–582.
Lewis, N.M., Chu, V.T., Ye, D., Conners, E.E., Gharpure, R., Laws, R.L., 2020. Household transmission of SARS-CoV-2 in the United States. Clin. Infect. Dis.
Li, F., Li, Y.Y., Liu, M.J., Fang, L.Q., Dean, N.E., Wong, G.W.K., et al., 2021. Household transmission of SARS-CoV-2 and risk factors for susceptibility and infectivity in Wuhan: a retrospective observational study. Lancet Infect. Dis.
Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., 2020. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N. Engl. J. Med. 382 (13), 1199–1207.
Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., 2020. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science 368 (6490), 489–493.
Li, W., Zhang, B., Lu, J., Liu, S., Chang, Z., Peng, C., et al., 2020. Characteristics of household transmission of COVID-19. Clin. Infect. Dis. 71 (8), 1943–1946.
Lin, J., Yan, K., Zhang, J., Cai, T., Zheng, J., 2020. A super-spreader of COVID-19 in Ningbo city in China. J. Infect. Public Health 13 (7), 935–937.
Liu, Y., Eggo, R.M., Kucharski, A.J., 2020. Secondary attack rate and superspreading events for SARS-CoV-2. Lancet 395 (10227), e47.
Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E., Getz, W.M., 2005. Superspreading and the effect of individual variation on disease emergence. Nature 438 (7066), 355–359.
Luo, L., Liu, D., Liao, X., Wu, X., Jing, Q., Zheng, J., 2020. Contact settings and risk for transmission in 3410 close contacts of patients with COVID-19 in Guangzhou, China. Ann. Intern. Med. 173 (11), 879–887.
Madewell, Z.J., Yang, Y., Longini Jr., I.M., Halloran, M.E., Dean, N.E., 2020. Household transmission of SARS-CoV-2: a systematic review and meta-analysis. JAMA Netw. Open 3 (12), e2031756.
O'Driscoll, M., Ribeiro Dos Santos, G., Wang, L., Cummings, D.A.T., Azman, A.S., Paireau, J., et al., 2020. Age-specific mortality and immunity patterns of SARS-CoV-2. Nature.
Shim, E., Tariq, A., Choi, W., Lee, Y., Chowell, G., 2020. Transmission potential and severity of COVID-19 in South Korea. Int. J. Infect. Dis. 93, 339–344.
Sun, K., Wang, W., Gao, L., Wang, Y., Luo, K., Ren, L., et al., 2020. Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. Science.
Tong, Z.D., Tang, A., Li, K.F., Li, P., Wang, H.L., Yi, J.P., et al., 2020. Potential presymptomatic transmission of SARS-CoV-2, Zhejiang Province, China, 2020. Emerg. Infect. Dis. 26 (5), 1052–1054.
Tsang, T.K., Wu, P., Lin, Y., Lau, E.H.Y., Leung, G.M., Cowling, B.J., 2020. Effect of changing case definitions for COVID-19 on the epidemic curve and transmission parameters in mainland China: a modelling study. Lancet Public Health 5 (5), e289–e296.
Yang, Y., Longini Jr., I.M., Halloran, M.E., 2006. Design and evaluation of prophylactic interventions using infectious disease incidence data from close contact groups. J. R. Stat. Soc. Ser. C Appl. Stat. 55 (3), 317–330.